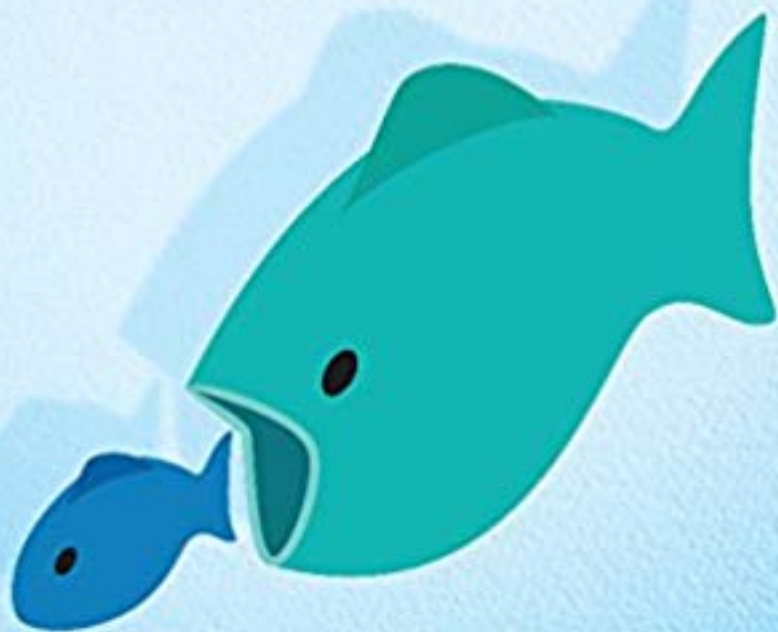


Exclusionary Practices

The Economics of Monopolisation
and Abuse of Dominance

Chiara Fumagalli, Massimo Motta
and Claudio Calcagno



Introduction

Exclusionary practices are contracts, pricing strategies and more generally actions taken by dominant firms to deter new competitors from entering an industry, to oblige rivals to exit, to confine them to market niches, or to prevent them from expanding, and which ultimately cause consumer harm. This is certainly the most controversial area in competition policy, and one in which economics has arguably not yet been able to guide policymakers in the design of sensible rules and enforcement practices.

Whether due to the influence of the Chicago School (in whose teaching there is little room for the possibility that dominant firms exclude rivals in a welfare-detrimental way) or due to other reasons (such as the expectation that entry will take place, hence reducing any existing market power), it is rare for US courts to find that a firm has infringed antitrust laws on the basis of monopolisation or attempted monopolisation.¹ In general, therefore, even firms with very significant market power are free to engage in unilateral business practices such as tying, exclusive dealing contracts, fidelity discounts and aggressive price policies (obviously, this lenient stance does not extend to coordinated behaviour such as cartels, which is punished very severely).

¹ Administrability may also have contributed to a more laissez-faire approach in dealing with exclusionary practices in the US. Indeed, Kovacic (2007) argues that it is the combined effect of the Chicago School (stressing that it was unlikely that certain practices would be anti-competitive) and of the Harvard School (calling for simple rules in order to make competition law easy to administer) that has led to a conservative stance in monopolisation cases. Note, however, that administrability may equally support simple rules in the other direction. In the EU, for instance, one often hears voices calling for blanket prohibitions of practices such as exclusive dealing and loyalty rebates since a more nuanced approach would be too complex for lawyers and judges to administer.

At the other extreme, dominant firms in the European Union ('EU') are under close scrutiny,² and it is very unlikely that cases involving practices such as exclusive dealing, fidelity rebates and price discrimination are decided in favour of a dominant firm.³

Most economists have denounced this state of affairs as unsatisfactory for quite some time⁴ and have emphasised that these practices may be anti-competitive or efficiency-enhancing depending on the circumstances. As a consequence, they should be neither under a (*de jure* or *de facto*) *per se* illegality nor under a *laissez-faire* regime, but should be assessed on the basis of the effects exerted on the market. Admittedly though, the guidance that economic theory has so far been able to provide to competition law enforcement in this area is not fully adequate. Some so-called post-Chicago models have offered what economists call 'possibility results' (namely, the development of models showing that a given practice *may* have an anti-competitive effect under certain conditions), but few 'general identification' results, which could assist the analyst in uncovering all the potential effects (positive and negative) of an exclusionary practice, as well as their significance in practice. Note also that such issues are extremely important for a modern economy, because wrong policies in this area can have welfare-detrimental effects either by eliminating competition (a hands-off approach would allow incumbent firms to exclude new or small efficient rivals, thereby leading to persistent dominant positions) or, at the other extreme, by impeding practices which lead to lower prices or higher investments (think of interventionist policies which prevent dominant firms from offering good deals, or from introducing new products, or from using contracts which may promote investments).

The objective of this book is to deal precisely with these issues, by developing a general analytical framework which encompasses and extends previous works, and by identifying clear and workable criteria that can help competition authorities in dealing with exclusionary practices. Indeed, an economics-based approach need not be a case-by-case approach, and it is important to find workable rules which allow competition authorities and

² Under EU law, a dominant firm has a special responsibility not to allow its behaviour to impair genuine, undistorted competition on the internal market.

³ The European approach has certainly been influenced by the so-called 'ordo-liberalism', a doctrine developed in Germany in the first half of the twentieth century, and according to which the law should protect the market from both ('unfair') distortions by public power (government) and by private economic power (large firms). See Amato (1997) for a discussion of the origin of competition law in both the US and the EU.

⁴ See in particular Vickers (2007).

courts to administer the law in a clear and predictable way, consistently with the principle of legal certainty.

Exclusion: a general analytical framework

There is by now a large body of economic models providing examples of why and how a dominant firm may exclude rivals in a welfare-detrimental way. In this book, we show that many of these models of exclusionary conduct are particular instances of a common mechanism, which hinges on the existence of scale economies and incumbency advantages (that is, an asymmetry between the dominant firm and the rival(s), for example in terms of established customer base, investment in a key infrastructure, exploitation of scale economies) that are found in a variety of industries. Where scale economies exist,⁵ a firm intending to challenge the dominant incumbent firm needs to attain a certain scale to be profitable. In turn, this means that if the dominant firm induces enough buyers to buy from it, the entrant will be deprived of the scale it needs and will refrain from entering or from expanding its operations beyond some market niche, or it will be obliged to cease operations. This will leave the dominant firm free to exercise monopoly power upon the remaining buyers and to recoup the loss (if any) it may have incurred while attracting the critical mass of customers away from the entrant. In this framework, there are different ways in which a dominant firm may attract buyers, for example: pricing below costs to some early buyers or markets, or to some large customers, engaging in exclusive dealing contracts with customers, tying a monopolised good to another good produced also by rivals, and refusing to deal with a competitor. By adopting this general framework, the book analyses a number of practices which may lead to exclusionary effects, and identifies under which conditions these practices are likelier to generate anti-competitive effects. This book shall also deal with exclusionary models other than the above-mentioned mechanism based on scale economies, with the aim of offering a more complete treatment of exclusionary practices; as well as policy implications which are sufficiently general and well grounded in order to provide some guidance to competition authorities and courts.

⁵ As will be emphasised throughout the book, such economies of scale may arise on the supply-side, for instance due to fixed costs or minimum efficient scale of production; or on the demand-side, for instance due to externalities among users such as in network and two-sided markets.

Specifically, this book analyses different practices across five chapters: Chapter 1 examines predatory pricing; Chapter 2 selective discounts (rebates) and other forms of price discrimination; Chapter 3 exclusive dealing; Chapter 4 tying and bundling; and Chapter 5 practices which may lead to vertical foreclosure, such as refusal to deal, denial of interoperability and margin squeeze.

Some policy considerations

Possession *versus* abuse of a dominant position

Competition laws in most jurisdictions do not prevent firms from obtaining or possessing a dominant position;⁶ what they do prohibit is that a dominant firm abuses its market position by preventing rivals from contesting its dominant position, thus hindering the good functioning of a market. This is notably the legal approach in the EU (see Article 102 of the Treaty of Lisbon on the Functioning of the European Union) and in the US (Section 2 of the Sherman Act),⁷ which have influenced most of the competition laws around the world. The principle that obtaining or possessing a dominant position is not by itself a problem is very important: it reflects the idea that it is the prospect of earning profits and market power which represents the engine of innovation and growth. Firms will innovate, invest, introduce new and higher quality products to be better

⁶ In the EU, the Court of Justice defined dominance as a ‘position of economic strength [...] giving [a firm] the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers’ (*United Brands*, para. 65). This legal definition relates to the economic concept of market power (the ability to set prices above marginal costs), in that the case-law will find dominance whenever the firm at issue enjoys substantial market power.

⁷ To be precise, the US Sherman Act does not use the ‘abuse of dominance’ terminology, in that it states: ‘Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony’. However, in practice, like in the EU, for a potentially exclusionary conduct to be found a violation of US antitrust law, there must be in addition to the possession of monopoly power (which by and large is a similar concept to dominance) ‘the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident’, as noted by the Supreme Court (see *Grinnell*, para. 570–71).

The main difference between the two jurisdictions is that the EU also condemns exploitative abuses (which may be thought of as unfair ways to exercise a dominant position), such as excessive prices, whereas the US does not. This book does not deal with exploitative abuses, but only with exclusionary ones. On the former, the interested reader may refer to Motta and de Streel (2007).

than rivals, be preferred by customers and hence earn higher profits.⁸ If in this process there is a firm which is doing so much better than the rivals that it will dominate the market, that should be accepted (some would say welcomed) – so long as there has been ‘competition on the merits’ and the firm has not resorted to unlawful means.⁹

If competition laws followed a different approach and found it illegal to *hold* a dominant position, then the competitive process would not work properly: knowing that it will not be allowed to earn high profits, a firm would have significantly weaker (if any) incentives to invest, innovate or introduce new products or new business models. In turn, customers (and final consumers) would not be able to enjoy new and better products or benefit from innovations, and the whole economy would suffer from lower efficiency levels. The principle that the firm’s incentives (that is, the prospect of earning high profits) should be preserved will be behind most of our policy discussions on how to treat certain practices. For instance, we shall argue in Chapter 5 that competition authorities should impose mandatory access to an input (for example, a technology or an infrastructure) belonging to a dominant firm only under exceptional circumstances.

The trade-off between intervening too much and too little

While there is probably consensus around the world that competition laws should not be designed or enforced to sanction the possession of a dominant position, but only its abuse, what actually determines an abuse and how competition authorities and courts should identify it, are clearly the most crucial and debated questions – and the main topic of this book. From the point of view of the case-law, it is fair to say – as we mentioned at the beginning of these introductory pages – that an infringement finding will be more likely in the EU than in the US, all else equal. In other words, in the EU (and in some jurisdictions around the world that have modelled

⁸ See also the US Supreme Court in *Trinko*: ‘The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices – at least for a short period – is what attracts “business acumen” in the first place; it induces risk taking that produces innovation and economic growth. To safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive conduct.’ (Part III of the Opinion)

⁹ In practice, the distinction between fierce – but fair or lawful – competition and unfair or unlawful competition can be difficult to make, as will be evident throughout the book (see also below in this introduction). One of our key objectives in this book is to provide guidance on how to make this distinction.

their competition laws after the EU version) there are more restrictions on the practices that a dominant firm can engage into than in the US. In the same vein, we feel that in the EU there has been, at times, the temptation to protect competitors rather than the competitive process, whereas in the US there is probably too much faith in the ability of the market to heal itself. Let us explain why we think both extremes are mistaken.

Empirical evidence shows that competition promotes efficiency and productivity growth mainly through a selection process.¹⁰ Absent rivalry, for instance in industries characterised by legal monopolies or where collusion is allowed, most firms would survive even if they are inefficient. When firms have to compete, instead, it will be those with good business ideas, which are well run and which continuously invest and improve their products and services, which will be successful and will grow. Whereas the least efficient ones – those which are badly managed, do not want to risk their capital, or quite simply have less appealing ideas or products – will have to downsize and might eventually have to shut down. This Darwinian process is the main source of productivity gains in an economy. But for this selection process to work, the market has to work well, and both entry and exit must be viable. In this light, neither an approach, which aims at protecting competitors, nor a *laissez-faire* approach, would serve the public interest.

If a competition authority is too prone to defend the rivals of a dominant firm even when the latter is competing on the merits, economic efficiency will not be promoted. Exit of inefficient firms is and should be part of the normal competitive process, and only too often do we forget that protecting inefficient competitors will have repercussions: the most efficient firms will not be able to take full advantage of their innovations, investments or business ideas, or will be dissuaded from offering pro-competitive price cuts, thus hindering the competitive process. A corollary of this approach is that competition authorities and courts should avoid protecting inefficient firms. By way of examples, in Chapters 1 and 2 we suggest that – when investigating a dominant firm for alleged exclusionary prices – competition authorities should adopt a safe harbour and find in its favour whenever its price is above an appropriate cost measure unless exceptional circumstances arise.¹¹ Otherwise, there would

¹⁰ See, for example, the surveys by Bartelsman and Doms (2000) and Syverson (2011).

¹¹ See, for example, the US Court of Appeals' judgment in *Barry Wright* (a case concerned with the allegedly predatory nature of prices that were above costs): '[W]e believe that [...] above-cost price cuts are typically sustainable; that they are normally desirable (particularly in concentrated industries); that the "disciplinary cut" is difficult

be the risk of protecting rivals which are inefficient (that is, they are unable to meet the incumbent's prices just because they have too high costs), and of chilling competition. In Chapter 3, we stress that – although potentially anti-competitive – an exclusive dealing contract with a buyer may also in principle lead to efficiency gains, for example where it protects investments made by the dominant firm in the specific relationship with that buyer. In Chapter 4, we caution against treating tying under a *per se* prohibition, because many innovations take place precisely by tying two previously separate components or products into a single one. Tying may well harm competitors in the markets at issue, but as long as consumers are benefiting from a genuine innovation, it would be difficult to conclude that the practice is anti-competitive.

On the other hand, while we believe that markets tend to function reasonably well, we should recognise that they can be (to a larger or smaller degree) imperfect: for example because of industry features such as very large fixed costs, sunk costs, switching costs or network effects; because of government regulations which raise legal barriers to entry; because of imperfect financial markets which make it difficult for young firms and potential entrants to obtain funds for a potentially good project. As a result, it may not be easy for rivals to compete effectively and contest the market position that a dominant firm has obtained in the past. In some circumstances, the market may not function well even absent particular strategic conduct by a dominant firm. If so, in some cases the correct response to a market failure may be the establishment of a regulatory regime. In other cases, a competition authority should remain vigilant and ensure that there is no conduct that the dominant firm resorts to in order to exclude rivals anti-competitively and cause harm to final consumers. This is because even practices which may appear to have limited effects if carried out by a non-dominant firm may actually have a significant impact if undertaken by a dominant firm. This appears to be the principle underpinning the existence of abuse of dominance provisions. As economists, we would add – as we explain throughout this book with reference to economic models – that the *degree* of dominance often goes

to distinguish in practice; that it, in any event, primarily injures only higher cost competitors; that its presence may well be “wrongly” asserted in a host of cases involving legitimate competition; and that to allow its assertion threatens to “chill” highly desirable procompetitive price cutting. For these reasons, we believe that a precedent allowing this type of attack on prices [...] would more likely interfere with the procompetitive aims of the antitrust laws than further them’ (para. 30).

hand in hand with the potential exclusionary effect of a given practice, all else equal.¹²

Different policy rules according to different conditions
across jurisdictions

As just set out, striking the right balance between over-enforcement (protecting inefficient competitors) and under-enforcement (letting a dominant firm take unfair advantage of its market position, causing in turn consumer harm) is not straightforward. Throughout the book, we suggest some policy rules which recognise and seek to resolve this trade-off in a reasonably effective way.

We need to acknowledge, however, that there may be economic, historical, institutional and legal considerations which may affect the optimal policy rules, particularly in the area of exclusionary practices. For example, the fact that the US approach has been typically less interventionist than the one in the EU may well reflect different economic contexts. In the US, where markets have generally been open and entry has tended to be relatively easy (because of potentially lower administrative barriers and because in a larger fully integrated market fixed entry costs may be recovered more easily), it may be safer to rely on market forces to solve exclusionary issues. Less so in Europe, which has traditionally known persistent positions of market power and less dynamic markets, or in less developed countries. Tapia and Roberts (2015), for instance, state that in developing countries ‘entrenched interests [...] have cornered certain markets and the rents that can be earned’. They associate this with various factors, such as large economies of scale (relative to the size of the local markets), obstacles to transport, the influence of well-connected business groups and families and the legacy of state support. For all these reasons, it is less likely that such countries may rely upon market forces to the same extent as in the US.

As a consequence, certain policy rules which may be relatively uncontroversial in the US may perhaps not fit less developed economies equally well. In the context of this book, in Chapter 5, for instance, we suggest

¹² A peculiarity of the approach to the enforcement of Article 102 is that the European Commission and the Courts seem ready to accept that certain practices (for example, exclusive dealing) may be legitimate when adopted by a firm which is not dominant; but as soon as the dominance threshold is met, the same practice may be presumed anti-competitive, potentially even at low levels of dominance. From an economic perspective, dominance (that is, market power) does matter, but it is a question of *degree* rather than a *binary* concept.

several limiting principles for intervention in refusal to deal cases, such as indispensability of the input and scarce investment on the part of the owner of the input. This is because in striking the balance between, on the one hand, protection and incentivisation of investments and, on the other hand, promotion of competition, it is the former which should typically be privileged. But in a less developed economy, where entry is difficult and rare, and if the input is owned by an entrenched ‘super-dominant’ firm which has historically enjoyed a privileged position, a more interventionist approach by competition authorities might be justified.¹³

Effects-based *versus* form-based approach

In the mid-2000s, the two US competition authorities (Department of Justice and Federal Trade Commission) made an effort to draw guidelines on how to enforce Section 2 of the Sherman Act.¹⁴ Building on a number of hearings of academics and practitioners, the ensuing Section 2 report and recommendations were adopted by the Department of Justice in 2008, while a majority of the Federal Trade Commission’s panel of commissioners opposed them, mainly on the ground that they would have led to too weak enforcement of anti-competitive conduct by single firms. This report was eventually withdrawn in May 2009.¹⁵ More recently, courts and commentators alike have hotly debated how to deal with pricing conduct and the circumstances under which above-cost pricing should constitute a safe harbour (see our discussion of *Meritor* and *Eisai* in Chapter 2).

In the EU, most of the discussion on how to enforce abuse of dominance provisions has revolved around whether or not to adopt an *effects-based* approach rather than a *form-based* approach. The former assesses practices by the effects they have on consumer welfare, independently of the form

¹³ In the same vein, Motta and de Streel (2007) argue that excessive pricing actions in antitrust may be justified in situations where (i) there are high and non-transitory barriers to entry and a very strong dominant position unlikely to be challenged, and (ii) the super-dominant position is the result of exclusive or special rights and legal concessions. Both conditions are less likely to be satisfied in the US than in both Europe and (especially) some developing countries.

¹⁴ See <http://govinfo.library.unt.edu/amc/>, setting out the role, activities and outputs of the (now defunct) Antitrust Modernization Commission.

¹⁵ For further background, we direct the reader to a speech by the then Federal Trade Commissioner J. Thomas Rosch (‘Thoughts on the Withdrawal of the DOJ Section 2 Report’, delivered to the IBA/ABA Conference on Antitrust in a Global Economy on 25 June 2009); and to a speech by the then Assistant Attorney General within the Antitrust Division at the Department of Justice Christine A. Varney (‘Vigorous Antitrust Enforcement In This Challenging Era’, delivered to the United States Chamber of Commerce on 12 May 2009).

they take. Consider, for instance, an agreement between a buyer and a supplier which commits the buyer not to purchase from rival suppliers (we refer to this as exclusive dealing throughout the book). Under a form-based approach, exclusive dealing by a dominant firm would be typically found anti-competitive by its very nature, independently of the market circumstances and of their effects. However, exclusive dealing may also protect investments by the supplier, and economic principles would call for an assessment of the ultimate effects on consumers, that is, whether pro-competitive effects balance (or even outweigh) anti-competitive ones or not. Accordingly, an effects-based approach would treat different practices having the same effects in the same way. By contrast, a form-based approach may end up treating them differently precisely because the practices take a different form, and hence fall into different categories. But this may have the perverse consequence of promoting ‘abuse-shopping’ by dominant firms, namely looking for the practice which – while achieving the same exclusionary objective – has the highest probability of staying below the radar of a competition authority or court and thus remain unchallenged.

In the mid-2000s, the European Commission started to reconsider the form-based approach followed (and fully endorsed by the courts) until then.¹⁶ This resulted in the publication of a Discussion Paper on the application of Article 102 to exclusionary practices,¹⁷ and in the Commission issuing a Guidance Paper on the enforcement priorities¹⁸ which were fully aligned with economic thinking and announcing an effects-based approach in the Commission’s forthcoming enforcement in this area.

As we discuss in Chapter 2, it is not clear to what extent the Commission has followed in practice such effects-based principles. More importantly, several court judgments appear not to have moved from a form-based approach. This perhaps has culminated with the *Intel* judgment of the

¹⁶ A role was perhaps played by the *Michelin II* case, that many perceived as having gone too far in this formalistic approach to Article 102. Prior to that judgment, standardised volume discounts – unlike most rebate schemes – had been deemed legal, but the General Court considered that they were anti-competitive as well, despite possible efficiency justifications and despite the existence of strong competitors (whose market share was increasing even during the period of the alleged infringement) – see Motta (2009) for an account of this case. It may also matter that the European Commission had already embraced an economic approach when dealing with mergers, so that an alignment of enforcement may have appeared necessary.

¹⁷ Directorate-General for Competition (European Commission) (2005).

¹⁸ European Commission (2009).

General Court, which established that rebates granted upon the condition that the customer makes most or all of its purchases from a dominant firm are anti-competitive by their own nature, independently of market coverage, market circumstances, or any other consideration. At the time of writing, the case is still pending on appeal at the Court of Justice, and it is unclear whether this will follow the General Court or the opinion of the Advocate General, who expressed strong criticism of the General Court's judgment, in what almost appears as a manifesto of an effects-based approach to Article 102. Either way, the *Intel* case shows that this is a debate which is probably going to last.

In this debate, we obviously stand by the idea that practices should be assessed according to their effects rather than by their form, and this book aims to offer guidance in this perspective. In particular, our objective is to identify under what circumstances certain practices are more or less likely to be anti-competitive, what factors one should look at in order to identify possible anti-competitive effects, and what are the theories of harm (or the likely pro-competitive rationales) which may underpin a certain conduct.

Some practitioners claim that an effects-based approach would entail long and complex cases well beyond the administering ability of judges, resulting in unpredictable outcomes and little legal certainty for firms. We hope that the discussion in this book will (at least partially) dispel this view. While an effects-based approach may involve more questions and analyses than a form-based (or a *per se*) approach, we believe that economics can offer guidance as to which questions and analyses one ought to focus on. We also note that judges in several jurisdictions (for example, US, UK, South Africa) do take economic considerations (including likely and actual effects of a conduct) into account in the enforcement of competition law in the area of exclusionary practices; and in other areas – most notably merger control - judges fully engage with very sophisticated economic arguments in many jurisdictions, including in the EU. In sum, if followed effectively, an effects-based approach would not come at the expense of legal certainty or of administrability.

How to use this book

Each chapter is structured broadly as follows. First, we introduce the range of practices considered in the chapter. Next, we present the main theories of harm underpinning the potentially anti-competitive effects of a certain practice; we do so with reference to the academic economic literature and to novel models, but limiting technical jargon and mathematical notation to

a minimum. We then, separately, examine in a formal (mathematical) way the economic models previously introduced. Depending on the nature of the practice, possible pro-competitive effects (or efficiencies) are discussed along anti-competitive effects or together with the policy implications that one may draw from the economic models presented in each chapter, in sections called *'From theory to practice'*. We then provide a short review of the key case-law¹⁹ in the US and, separately, in the EU. Finally, we review a few cases (from various jurisdictions, including beyond the US and the EU) in some detail; we have selected these cases because we believe that certain aspects (as presented in publicly available documents, such as decisions and judgments) may allow one to interpret some of them in the spirit of some of the economic models presented and may therefore provide helpful illustrations of the concepts discussed.

We have written this book with a diverse audience in mind, ranging from economics students and researchers to competition policy practitioners who are not economists. We use economic analysis throughout the book and our main conclusions are presented through formal models. The technical sections of the book (identifiable by shaded pages and headings marked with *) can therefore be used, for example, to complement a graduate or advanced undergraduate course in competition policy or industrial organisation. Trained economists should be able to capture the main intuitions of our analysis by reading the non-formal discussion of the economic models presented in each chapter, but may also want to go more carefully through some of the models set out in the technical sections, and possibly also want to follow up with the ample economic literature referred to. Practitioners who are not formally trained economists, instead, should be in a position to navigate through – and hopefully gain some insights from – any section other than the technical ones.

Notes on references to EU law

Throughout the book, references to the old numbering of the Treaty of Amsterdam and the Treaty of Rome have been updated using the numbering introduced by the Treaty of Lisbon on the Functioning of the European Union ('TFEU'). Hence, early references to Articles 85 and 81 are

¹⁹ In these 'case-law' sections, we also include key decisions by competition authorities (in particular by the European Commission). While we are aware that such decisions do not technically constitute case-law, we think that a brief review of some of these may facilitate the discussion in these sections.

replaced by references to Article 101, and early references to Articles 86 and 82 are replaced by references to Article 102.

Further, following the renaming of the EU courts in the Treaty of Lisbon, we refer to the General Court, which was formerly known as the Court of First Instance, or ‘CFI’; and to the Court of Justice of the EU (or in short, Court of Justice), which was formerly known as the European Court of Justice, or ‘ECJ’.

Predatory Pricing

1.1 Introduction

Predatory pricing is one of the most interesting and most controversial issues in antitrust. The term refers to a practice whereby an incumbent firm (the predator) sets prices very aggressively with the aim of excluding a rival from the market (that is, forcing the rival to leave the market or discouraging it from entering) or marginalising the rival and relegating it to a niche role. Predation – if successful – will therefore be associated with the existence of two periods: one, the predatory period, in which consumers will enjoy low prices and the incumbent will sacrifice profits; the other, the recoupment period, in which the incumbent will be able to increase its prices, and obtain higher profits, because the prey is no longer in the market (or has been marginalised). From the incumbent's point of view, this strategy is profitable if the earlier profit sacrifice is outweighed by the subsequent higher gains. From the point of view of consumers, and of social welfare, exactly the opposite happens. If the predatory strategy is successful, higher surplus during the predatory period will be outweighed by lower surplus in the recoupment phase.

Given that predatory episodes are associated with low prices, it should not come as a surprise that it is extremely difficult to distinguish low prices that are an expression of tough but fair competition from low prices that are an expression of an exclusionary strategy by the dominant firm. Suppose we observe that after the entry of a competitor, an incumbent firm reacts by starting to cut prices aggressively. Is this the sort of genuine competitive response that we should expect (after all, any theory in which firms do not collude would foresee that entry would lower equilibrium prices), or is it instead predation? In other words, are low prices good news for consumer welfare, or are they instead just a temporary consumer

gain which anticipates the permanent consumer loss which will take place when the prey exits the market?¹ This is the first reason why predation is a controversial issue. A second reason is that, until the 1980s, economists had been unable to propose a convincing rationale for predation. As we shall explain in Section 1.2.1, Chicago scholars argued that predation is not a profitable and rational strategy for a dominant firm, and therefore that we should not expect predatory episodes to occur at all. This view was – and probably still is – quite influential and is reflected in the current status of the case-law in the US, where plaintiffs have not been successful in a predatory case for a long time (see Section 1.5). Since the 1980s, modern industrial organisation theory has been able to find rigorous explanations of why an incumbent firm may have an incentive to use predatory prices to exclude a more efficient rival.² These theories – which we shall call ‘traditional theories’ – are mostly based on the existence of imperfect information³ and will be reviewed in Section 1.2. The focus of this chapter, however, will be on an alternative theory which does not rely on imperfect information, and which is based instead on the co-existence of scale economies, an incumbency advantage and sequential buyers (or markets) (see Sections 1.2.3 and 1.3). Note that this theory (based on Fumagalli and Motta, 2013) is new, in the sense that it rationalises predation in particular cases where traditional theories may not apply, but does not replace traditional theories. Indeed, the mechanism underpinning this new theory might well co-exist with other rationales for predation.

In Section 1.4, we will comment upon the implications of such theory for competition policy and the features that an industry must exhibit for such predation theory of harm to hold. This may provide a guide to distinguish unlawful predatory pricing from lawful price-cutting.

In Section 1.5, we will give an overview of landmark decisions by competition authorities and of key case-law in the US and in the EU, while

¹ Note, incidentally, that one would like to distinguish unlawful predation from lawful price-cutting when the predatory episode is still occurring, before the prey might be obliged to exit the market. However, even *ex post*, the identification of predatory pricing is not easy: observing that the incumbent increases its prices after exit is not sufficient to prove predation, because – once again – any model of competition would tell us that prices will increase when the number of competitors is reduced, all else equal.

² The discussion has moved to the extent to which these theories provide workable criteria to identify real cases of predation.

³ We note that, in economic theory, the notion of imperfect information encompasses that of incomplete information and of asymmetric information. In this chapter, for simplicity, we will generally refer to ‘imperfect information’, although we note that some of the models we refer to considered specific instances of it.

in Section 1.6 we will review a number of international cases on predation which can be read in the spirit of the economic models we present in this chapter.

1.2 Theories of Predation

In this section, we briefly describe how the economic literature has dealt with predation so far.⁴ First, we report and discuss the Chicago School's skeptic views about predation, as expressed by McGee (1958). Next, we summarise the traditional economic models which offer a rationale of predation on the grounds of imperfect information. Finally, we discuss a theory of predation based, instead, on the existence of scale economies and an incumbency advantage (originally due to Fumagalli and Motta, 2013).

1.2.1 Predation: The Chicago School and the Search for a Theory

Although cases of predatory pricing were not rare in the US, at least until the *Brooke* ruling in 1993, theories which could formally (that is, using rigorous economic modelling) explain predation did not appear until the 1980s. For a long time, the main (informal) explanation for predation was probably the 'deep pocket' (also called 'long purse') story, which goes roughly as follows. A large firm might drive a small competitor out of the market by waging a price war that generates losses to both. But the small competitor has limited resources (a 'small pocket') and will therefore be unable to survive such losses for a long time. Sooner or later, it will have to give up and leave the industry, allowing the large firm to increase prices and recoup losses. Unfortunately, however, a solid theory to support this story has appeared only relatively recently (see below).

The idea that a firm could drive out competitors by using predatory pricing was criticised on four main grounds by McGee (1958), in a very influential article. First, due to its larger market share, a large firm will usually suffer greater losses from a price war than a small firm: other things being equal, the same unit loss will be multiplied by a larger number of units. Second, predation makes sense only if the large firm will increase prices when the prey leaves. But, McGee argues, the assets and plants of the small firm will not disappear, and as soon as prices rise the small firm can re-enter, or its assets might be bought and used by somebody else, reducing the profits the predator can expect to make. Third, the predation theory assumes that the predator has a deep pocket and the victim a small one,

⁴ This section partly follows Motta (2004: Section 7.2).

while this should rather be explained than assumed. In this perspective, one should wonder why a small firm, even if financially constrained, could not set out the situation (including the fact that the predator is making more losses than the small firm does, and cannot sustain them forever) to its creditors, thereby obtaining funds until predation ends. Fourth, for predation to be rational, it must be not only feasible but also more profitable than alternative instruments. If a large firm wanted to get rid of a competitor, this criticism goes, predation would be an inefficient tool because it destroys industry profits for the time it lasts. Taking over the rival would be a more profitable strategy, as it would allow the preservation of high profits in the industry.

The first two arguments above can be taken care of relatively easily. Indeed, the first point does not hold if the large firm could price discriminate and decrease prices selectively only in those markets or for those clients where the small firm is competing. This allows the predator to preserve high margins on most of the units it sells, therefore reducing the cost of the predation strategy.

As for the second point, it relies on the idea that entering and re-entering the industry does not entail sunk costs. But fixed sunk costs are pervasive, and a firm cannot typically close down its plants, fire its workers, cease to supply its product one day and return costlessly to its business soon after that.

Furthermore, the very fact that the incumbent has successfully preyed once may have an influence over other firms that are considering entry into the same market. A potential entrant will not rush into the industry after seeing the end of its predecessor. This is one of the important counter-objections made by Yamey (1972), who pointed out that predation will discourage further entry into the industry. If an incumbent develops a reputation for reacting toughly and aggressively towards entry, potential competitors might be discouraged from entering the industry at all. Although it has taken game theorists some time to prove this reputation argument formally, it is now rigorously established, as we shall discuss in Section 1.2.2.2.

Perhaps the most challenging point made by McGee is the third. Suppose that the incumbent is indeed endowed with more financial resources than a small rival, although they are equally efficient. Why should the small firm not be able to get further financing from banks or other lending institutions to resist the predatory attack? After all, outside investors should understand that predation could not be successful if they gave unlimited funding to the prey. Anticipating that, predation would

not take place at all. Again, it is only recently, with the developments in corporate finance theory, that a convincing story has emerged of why predation might make the financial constraints of firms tighter, as we shall discuss in Section 1.2.2.1.

Finally, note that the fourth point made by McGee stresses an important general issue, namely that predation must not only be feasible but also more profitable than alternative options available to the incumbent. On the particular point that a takeover would be more profitable than predatory prices, three counter-objections can be made. First, buying out a competitor might encourage new ones to enter the industry with the aim of selling out to the incumbent at a profit: if it gains the reputation that new competitors will be bought out, a merger might not be a cheap option. Second, under some antitrust laws, taking rivals over might not be allowed for dominant firms. Third, as both Telser (1966) and Yamey (1972) argued, predation and mergers are not necessarily mutually exclusive options: aggressive price behaviour might well result in the prey being ready to sell out at a lower price (this would be the so-called 'predation-through-merger'). The merger strategy is therefore not necessarily in contradiction with a predation strategy but may be complementary to it.

Indeed, Burns (1986) looked at the expenditures made by American Tobacco to take over 43 competing firms between 1891 and 1906, and found econometric evidence that predation substantially decreased the acquisition prices. Aggressive price behaviour helped both directly (by reducing the price of acquiring a victim) and indirectly (by establishing a reputation for being a predator, that persuaded other rivals to sell out before any predatory episode would start).

This discussion of McGee's (1958) arguments and their possible counter-objections allowed us to point our attention to the main issues related to predatory pricing. In what follows, we briefly summarise some models which have addressed such issues.

1.2.2 Traditional Theories of Predatory Pricing

There is a common thread behind most of the traditional models of predatory pricing, namely the existence of some *imperfect information*. The predator will try to use the imperfect knowledge of the entrant (or of the outside investors that finance it), and behave so as to make them believe that the entrant would not make large profits in the industry. As a result, the entrant will exit, or will not enter, or its lenders will not be willing to provide it with more funding.

Three main types of imperfect information stories can be identified, based on: (i) financial market models of predation, (ii) reputation and (iii) signalling models. In what follows, we briefly summarise them. We shall also argue that financial market models are very similar in spirit to a theory of predation based on scale economies and incumbency advantages.

1.2.2.1 Predation in imperfect financial markets

As we have seen above, a weak point of the deep pocket theory of predation is that it does not explain why the prey has limited access to external funding. Modern corporate finance theory, focusing on the imperfections existing in capital markets, provides an answer to this question. This leads to a financial theory of predation (due to Bolton and Scharfstein, 1990) where the prey's limited access to external funding is endogenous, since predation affects the perceived risk of lending money, thereby reducing external financial sources available to the prey.

The key point of this theory is the existence of imperfect information on the side of the lenders (be they banks, equity holders, or other financial institutions). Lenders do not operate themselves in the borrower's industry and cannot have precise knowledge about it (or cannot observe some of the actions taken by firms). This characterises the relationship between the lender and the borrower. (In these principal-agent models, the bank is the 'principal', and the borrower the 'agent'.) Outside investors cannot be sure that the money lent is used in an efficient and competent way rather than being used by the entrepreneur for his or her private benefit, or in an exceedingly risky way (there is a so-called 'moral hazard' problem). Accordingly, outside investors will have to devise a contract that induces the borrower to 'behave', thereby protecting their interest in receiving the money back. For instance, suppliers of capital can require collateral, or can extend financing (in staged commitments) under a threat of termination in case of poor performance. Contracts of this type mitigate the agency problem, but generate another inefficiency in the form of *credit rationing*: if the borrower is not endowed with enough collateral, or if the initial performance has been inadequate, a project with a positive net present value might not be financed. In turn, this may invite predation in the product market.

Consider competition between an incumbent and a new firm. The incumbent is a well established firm that has accumulated enough financial resources in the past (liquidity, retained earnings, assets that can be used as collateral), whereas the new firm does not have enough own resources

and needs to borrow heavily to compete on a par with the incumbent. In such a situation, predation by the incumbent causes the prey's profits to fall, thereby limiting the amount of retained earnings (or more generally of collateral) the prey can put up to obtain external funds, or inducing the lender to deny the loan extension because of poor performance. It is therefore the aggressive behaviour by the incumbent that endogenously limits the access of the rival firm to external funds.

A possible objection to this predatory argument is that the lender, understanding that predation might occur, thus destroying its opportunity to make profits out of the loan, might have an interest in preventing it by announcing that it will finance the prey no matter its initial performance or the amount of collateral it owns. However, it might not be optimal for the lender to do so precisely because it would undermine protection from moral hazard (that is, inefficient behaviour by borrowers).

We believe this is a convincing story of why predation can take place. Aggressive market behaviour is used by the incumbent to modify the outside investors' evaluation of the risk of investing in the prey. As a result, the prey will have a lower ability to borrow and will be obliged to exit the industry or to reduce the scale of its operations.

Note that this theory does not predict that the incumbent has to set prices below costs to exclude the rival. The incumbent needs only to be sufficiently aggressive so as to make the rival's profits fall below the level required to obtain external funding. In doing that, the incumbent sacrifices profits but does not necessarily suffer losses.

Note also that this theory shares some similarity with the theory of predation based on scale economies that we will illustrate in Section 1.2.3. The basic mechanism is very similar: take away orders, or market, or profits, so as to make it impossible for the rival to survive – in the theory based on scale economies because this prevents the prey from achieving efficient scale, while in Bolton and Scharfstein's model because this denies access to external financing.

Empirical evidence We next summarise a few academic studies that provide empirical evidence consistent with the financial theory of predation.

Scott Morton (1997) looks at the British shipping cartel active at the turn of the twentieth century (that is, well before Europe had antitrust laws). Her econometric analysis shows that entrants were much likelier to be preyed against when they had fewer financial resources (as well as less experience, less multimarket contact with the cartel and a smaller customer base).

Lerner (1995) looks at the rigid disk drives industry of the 1980s. He documents that, in the earlier part of the period considered, when there was plentiful equity financing, there was no significant price difference among the products. However, in the latter part of the period studied (when financing became harder), drives whose closest substitutes were sold by undiversified thinly capitalised specialist firms were almost 20 per cent cheaper than drives without such close substitutes. Put simply, firms believed that it paid to be aggressive when aware that their close rival was in dire financial conditions.⁵

In addition, some economists have identified and commented upon specific real-world cases that could also be interpreted in the spirit of financial market predation. Gabel and Rosenbaum (1995) give a thorough account of the Wisconsin Telephone (part of the Bell family) case. Following Bell's patent expiration in 1894, Dane County Telephone began offering local services at half the price of the incumbent and signed up a good number of subscribers, including many switching away from Wisconsin Telephone. The entrant had difficulty though in raising capital to build larger (toll) networks. The authors argue that the strategy adopted by the incumbent was then to compete aggressively against the independent operator in its local markets (it undercut their price by 50 per cent). Lacking early profitability, the entrant was thus prevented from accessing financing opportunities and therefore from expanding into other markets.⁶ The financial predation strategy was viable and profitable, according to the authors, because the incumbent operated in more markets than the entrant and because there were common costs (sunk investments) across these markets, where network externalities further exacerbated the handicap faced by the entrant.⁷

⁵ The author controlled for a wide number of product features and the results were robust to a number of alternative definitions of what a 'close substitute' was. He also rejected the reverse causality (that is, the inability to raise capital for a firm producing drives similar to those of an efficient producer).

⁶ Eventually (in 1908) Dane County was sold to Wisconsin Telephone for a low price. This case can therefore also be read in the spirit of predation-through-merger, which we alluded to in Section 1.2.1, and to which we will return in Section 1.2.2.3.

⁷ Looking at the same industry, but in a different context, Weiman and Levin (1994) suggest that the Southern Bell Telephone Company gained its monopolistic position through financial predation. In particular, the authors argue, 'Southern Bell effectively eliminated competition through a strategy of pricing below cost in response to entry, which deprived competitors of the cash flow required for expansion even if it failed to induce exit.' This was complemented by pre-emptive investments in capacity, the authors also noted.

Bolton et al. (2000) report on the cable TV industry in Sacramento (California) in the 1990s. The authors first stress the existence of high entry barriers (facilitating recoupment post-predation) and a strong incumbency advantage. As for financing issues, the authors point out that the entrant heavily relied on the personal wealth guarantees of two of the owners to obtain its original \$6m investment; any extra outside financing was offered conditional on good market performance at an early stage. The incumbent reacted with drastic price-cutting, until the entrant was forced to cease its activities after a few months of losses.⁸

1.2.2.2 Reputation models

The main idea behind the reputational theory of predation is that the incumbent's aggressive conduct might allow it to create a reputation of being strong and aggressive, so as to discourage further entry into the market in the future.

To understand how Kreps and Wilson (1982) model this reputation-based theory of predation, suppose that there is an incumbent monopolist that is active in a number of identical markets, where it has the same technology and products (for example, a 'chain-store'). In each of these markets, it faces a potential entrant. Entrants can enter one at a time, in the following way. In the first market, first the potential entrant decides whether to enter or not, and if entry occurs the incumbent decides whether to fight or accommodate it. Then this same game is repeated, one by one, for all the markets.

Suppose also that the entrants have some uncertainty about the incumbent. When the game starts entrants believe that with some (possibly very small) probability the incumbent is 'strong', rather than 'weak' (or call it 'normal' if you prefer), where a 'weak' incumbent is one that has costs as high as the entrants and that, if the game was played only once, would not fight entry, because it would find this unprofitable. Fighting amounts to setting a low price that causes losses to both the entrant and the weak incumbent. Instead, a 'strong' incumbent is a very efficient firm whose costs are so low that fighting entry, that is, setting low prices that cause losses to the rival, represents its optimal strategy even in the short run.

Clearly, a strong incumbent will always fight entry, but this will not be predation: simply, it is so efficient that the entrant cannot coexist successfully. The interesting finding, as proved by Kreps and Wilson (1982),

⁸ The authors report that discovery led to the finding of internal documents drafted by the incumbent that explicitly sought to assess the viability of the entrant's business, as well as the net personal worth of its two principals.

is another: a weak incumbent would exploit the entrants' uncertainty and would fight entry at the beginning of the game, to establish a reputation for being strong and thus discourage further entry. It would be only towards the last periods of the game that a weak incumbent will accommodate entry, as the closer to the end of the game the lower the expected gain from pretending to be strong. In general, in any period, the weak incumbent's decision to fight reinforces its reputation to be efficient, but involves the sacrifice of current profits in order to deter entry and earn higher future profits. At the beginning of the game, the future is far enough and the trade-off is in favour of fighting, whereas at the end of the game there is less to be gained from deterring further entry (in the limit, in the last period there is no future gain at all), and the trade-off is in favour of accommodating.⁹

Empirical Evidence Commentators have identified a number of real-world cases whose facts seem to be consistent with the above theory. (In Section 1.6.11, we will discuss reputational issues in *Standard Oil* and *American Tobacco*.) Empirically, it is not a straightforward exercise, as this entails identifying entry that has not occurred but that would have occurred absent the reputation for predation enjoyed by an incumbent (where reputation itself is not a readily observable, quantifiable measure).

Gabel (1994) went through the business records of the telephony industry for the period from 1894 to 1910, with a focus on the US Midwest, and concluded that AT&T's dominance was due to predatory pricing.¹⁰ The main channel through which predation took place, the author noted, was the reputation that AT&T built when scaring off independent companies that considered starting long-distance phone services. The multiplicity of geographic markets lent itself quite naturally to the interpretation of a reputation model here: when an incumbent signalled that it was 'strong' in one market, the reputation effect could be felt in several other markets too.

But there was also a financial predation motive, the author noted: AT&T would force the independents to make losses in their local markets so as to make it more difficult for them to raise capital to build a long-distance network; and this exacerbated the financing constraints these firms were already facing.

⁹ Note that imperfect information is necessary for predation to exist. As showed by Selten (1975), if the entrants knew for sure that the incumbent is weak, they would never be 'tricked' into thinking that it is 'strong' instead. As a result, the incumbent would always accommodate entry and each entrant would enter the industry.

¹⁰ The author notes that none of these AT&T cases was fully litigated.

Last, we note that – to our knowledge – a predatory theory of harm *purely* due to reputation effects has never been successful before the courts on either side of the Atlantic. On the other hand, there have been a number of cases where reputation effects were raised alongside other anti-competitive theories of harm, as we shall note in Section 1.6, where we discuss cases that can be interpreted in the light of a number of economic models we present in this chapter.

1.2.2.3 Signalling models

Signalling models of predation are based, like reputation models, on imperfect information. Again, the potential entrant does not know whether the incumbent is low cost (strong, or efficient) or high cost (weak, or inefficient), and the incumbent will try to exploit this uncertainty to deter entry. The first signalling model is due to Milgrom and Roberts (1982) and it can be roughly summarised as follows.

Before making its entry decisions, a potential entrant observes the price set by the incumbent when it is still a monopolist. If it was certain that the incumbent is high cost, entry would be profitable. If it was certain that the incumbent is low cost, entry would entail a loss. Instead, the entrant is only in a position to conjecture that the incumbent is low cost with some probability, and it can only revise this probability by observing the monopoly price of the incumbent (if it enters, instead, it will immediately learn whether the incumbent is high cost or low cost). In this context, it is clear that a high cost incumbent might want to mimic a low cost one, to try and deter entry. However, a low cost one would not like to be mistaken for a high cost one, because it would attract entry, which lowers its profits.

There are two possible equilibria in the game. In the first (called ‘separating equilibrium’), the low cost incumbent will set a price lower than its normal monopoly price in the first period (when it is the only active firm), and this price is so low that no high cost incumbents would like to set it, because it would involve too high losses. Since there is no scope for mimicking the low cost incumbent, the high cost one will instead choose its normal monopoly price. The entrant will immediately learn which incumbent it faces: if the price is low, it can only be the low cost one, and it will stay out. If it is high, it will face the high cost incumbent and will enter.¹¹

¹¹ Note that in this equilibrium one could say that there is predation, in that the low cost incumbent is acting ‘strategically’ and sacrifices current profits to deter entry and gain more in the future. But, interestingly, its behaviour does not hurt welfare. To see why,

In the second equilibrium (called ‘pooling equilibrium’), instead, there is no price at which the low cost incumbent can profitably sell and be distinguished from the high cost one. As a result, it will simply set its normal monopoly price and the high cost incumbent will imitate it in order to deter entry.¹² In this case, we observe predation by the high cost incumbent, which sets a lower price than it would otherwise set in the first period (even though, it is important to notice that price might be above or below the incumbent’s costs), but it can act as a monopolist in the second: it sacrifices current profits to increase the future ones. The impact on welfare is more likely to be negative in this case.^{13,14}

Predation-through-merger An extension of the signalling model above allows us to explain why predation might be used to lower the price of taking over rivals, a strategy that has been alluded to in Section 1.2.1 (and to which we shall return in Section 1.6.11, where we discuss *Standard Oil* and *American Tobacco* in more detail).

Saloner (1987) changes Milgrom and Roberts’ model slightly, to allow for the possibility that firms merge after the first period (also, in his model, entering when facing a low cost incumbent would not give rise to losses, but just lower profits). In this case, setting a lower price than would otherwise be optimal signals to the entrant whether it should expect to make high or low profits after entry. (This argument was also partly present, informally, in Yamey, 1972.) This conjecture thus instructs the entrant whether it should be willing to sell out to (or merge with) the incumbent at a high or

note that in a perfect information world the entrant facing the low cost incumbent would never enter (by assumption of the model), and consumers would have to pay the normal monopoly prices in both periods. In this equilibrium, instead, the low cost incumbent charges a much lower price than it would otherwise do, to signal its efficiency. Therefore, while in the second period the price is the same, consumers will be better off in the first period. In a sense, by signalling its true nature through low prices the low cost incumbent is providing a service that enhances social efficiency.

¹² The entrant does not learn anything from the observation of the first period prices, and decides on whether to enter or not on the basis of its *ex ante* probability of facing a weak incumbent. For the pooling equilibrium to exist, this probability must be low enough: the entrant will stay out only if it expects a high likelihood to face an efficient incumbent. If it expected with a high probability to meet a weak incumbent, it would enter. But then, it could not be an equilibrium as the high cost incumbent would have no reason to sacrifice current profits if it knows it will not deter entry.

¹³ To be precise, the net effect is ambiguous *a priori*, since it involves a gain in the first period and a loss in the second.

¹⁴ Note that signalling models of predation are not inherently associated with the incumbent setting a low price. If the entrant – new to the industry – does not really know its own costs, and expects them to be highly correlated with those of the incumbent, then the incumbent might deter entry by setting a *high* price, because this would signal the existence of high costs in the industry for both (see Harrington, 1986).

at a low price. Again, predation takes the form of setting lower prices than a short-run calculation would imply, but this time its objective is not to deter entry, but rather to improve the terms on which the rival will accept being taken over.

Other predation models There are several other models where the incumbent might want to act strategically so as to make the entrant (or an existing competitor) expect lower profitability if it entered (or if it stayed in) the industry. Scharfstein (1984), for instance, analyses a model of ‘test-market predation’, where the entrant has a new product and is uncertain about the demand for it. Given this uncertainty, it introduces the product in a test-market first to see how it would be received. The incumbent might engage in various predatory practices (for instance secret price discounts to consumers) to make the entrant believe demand for its product will be low, thus leading it to abandon the market or reduce its scale of activity.

Fudenberg and Tirole (1986) suggest that the incumbent might also engage in ‘signal-jamming predation’, not to allow the entrant to improve its information. In a test-market model, for instance, the purpose of the entrant is to gather information about demand, and the predator defeats this purpose by openly cutting prices. The entrant knows that its demand is artificially low due to the incumbent’s cut-throat prices, but it cannot have any information about what demand would be under normal competitive circumstances. In the absence of information, it will prefer to exit. A similar ‘signal-jamming’ mechanism might also be used in other circumstances where there is imperfect information.

1.2.3 A Theory of Predation Based on Scale Economies and Incumbency Advantage

We next propose a theory of predation based on the existence of scale economies – whether on the supply-side (unit costs of production decreasing with the number of units) or the demand side (such as when the utility derived from a product or service increases with the number of its users) – and an incumbency advantage.¹⁵ Suppose that at a given moment in time, a dominant firm has an initial (incumbency) advantage over a rival: the former enjoys greater scale economies than the latter, perhaps because it has more market outlets, more captive customers, or a larger installed base than the rival. Suppose also that – in order to be profitable – the rival firm needs to reach a certain scale, that is, a certain

¹⁵ See Fumagalli and Motta (2013) and Section 1.3 for a formal treatment.

number of clients, or a certain installed base. Imagine that the dominant and the rival firm compete for a set of new buyers or for new markets (or for buyers/markets which are contestable).

In an industry of this type, the incumbent may engage in aggressive pricing to some early buyers (or in markets which develop first) to deprive the rival of the scale it needs to operate successfully. Once deprived the rival of key buyers (or markets), the incumbent will be able to raise prices on the remaining later buyers (or markets which develop subsequently), thereby recouping losses. The two usual ingredients of predation, early sacrifice of profits followed by later recoupment, are therefore present in a theory of harm based on a scale economies and incumbency advantages as well.

It is worth noting that this theory – that we develop formally in Section 1.3 – does not rely on imperfect information (unlike most of the traditional models of predation reviewed in Section 1.2).¹⁶ It is the interaction between scale economies and an incumbency advantage which makes exclusion possible. Hence, the incumbent may exclude an as-efficient or even a more efficient rival even if the latter can approach buyers and submit bids at the same time as the incumbent.¹⁷ Finally, note that our claim is that the existence of scale economies and of an incumbency advantage *may* lead to predation, not that predation always takes place under such circumstances. We will discuss in Section 1.4.2 the factors that we have identified as crucial in the theory to provide some guidance to competition authorities when dealing with predation cases.

1.2.3.1 A simple example

Perhaps the simplest setting to see the mechanism just introduced at work is the following. Suppose that the incumbent and its (smaller) rival compete for two *new* (that is, contestable) customers, or two new cohorts of customers, each with an order of one unit. The incumbent has a constant marginal cost of production $c_I > 0$. The rival's cost is equal to f for the first

¹⁶ Also Cabral and Riordan (1994, 1997) and Farrell and Katz (2005) rationalise predation in the absence of information asymmetries. In Section 1.2.3.2, we will discuss the relationship between their theories and ours. Another model that does not rely on information asymmetries is Harrington (1989), where *joint* predation is used to sustain collusion when entry barriers are low: if entry took place collusive firms would implement a policy of predatory prices.

¹⁷ If the incumbent also enjoys a *first-mover advantage*, that is, if it can make offers to buyers before the rival could react and make counter-offers, exclusion is easier and does not necessarily require any sacrifice of profits by the incumbent. In our setting, instead, the incumbent needs to sacrifice profits on early buyers to achieve exclusion. Most of the literature on the anti-competitive effect of exclusive contracts, that we discuss in Chapter 3, assumes a first-mover advantage for the incumbent.

unit and to 0 for the second unit, with $c_I < f < 2c_I$. This cost structure depicts a situation in which the incumbent – possibly because it can rely on a larger base of captive customers than the rival – has already exploited scale economies and can supply the contestable buyers at a constant marginal cost. The rival, instead, needs to sell to *both* contestable buyers to achieve efficient scale. When that is the case, the rival produces at lower total costs than the incumbent, and for this reason it would be beneficial for society that the rival supplies the contestable buyers. However a *single* contestable buyer is insufficient for the rival to reach efficient scale and for this reason it suffers a cost-disadvantage vis-à-vis the incumbent when the supply of a single buyer is concerned. We will discuss below plausible situations in which such a cost structure may arise.

The incumbent and the rival can both make price offers to get the customers' orders. Suppose also that competition takes place sequentially, in the sense that first firms compete for the first buyer, and that after this buyer has decided from whom to buy, they will compete for the second buyer.

The cost structure described above implies that – if for some reason it were able to secure the first buyer – the incumbent would be able to extract higher revenues than the rival from the second buyer. The intuitive reason is that the rival is particularly inefficient in supplying a single contestable buyer: $f > c_I$ (that is, the cost of the rival of supplying one unit is higher than the cost of the incumbent). Hence, if the incumbent secures the first buyer, then it will face a very weak competitor in the second period, and it will be able to 'win' the second buyer by setting a relatively high price. If instead, it is the rival who secures the first buyer, the incumbent will not be such a weak competitor in the following period (its cost of supplying one unit is c_I , which is lower than the rival's cost of supplying the first unit, f), thereby limiting the revenues that the smaller rival can extract in the second period.¹⁸

¹⁸ The formal proof presented in Section 1.3 will clarify this statement. Intuitively, when a given supplier (say firm i) secures the first buyer, competition for the second buyer involves one firm (firm i) who has already produced for one customer and the opponent who has not. Scale economies imply that firm i , who has already supplied the first buyer, is more efficient in serving the second buyer, thereby winning the competition in the second period. The key point is that the price that firm i will charge to the second buyer (and the rents it will be able to extract) will be pinned down by the opponent's cost to supply a *single* customer. Now, compare the situation where it is the incumbent that has secured the first buyer to the situation where it is the rival that has managed to. The assumption that the rival is less efficient than the incumbent in supplying a *single* customer, implies that the rents extracted by the incumbent in the former case are larger than the ones extracted by the rival in the latter case.

When competing for the first buyer, each firm realises that whoever wins the first buyer will also win the second one, and this will result in aggressive bidding for the first buyer. Who will win the first buyer, then? There are two effects at work: on the one hand, the incumbent is less cost-efficient over the two units than the rival, and this limits its price aggressiveness; on the other hand, the incumbent expects to extract higher revenues than the rival from the second buyer, and this makes it more aggressive.

It is possible to show (for a formal treatment see Section 1.3.1) that if the (overall) efficiency advantage of the rival is not strong enough (that is, if c_I is sufficiently close to 0, other things being equal), then it is the incumbent which will make the winning bid for the first buyer. Therefore, predation will arise at the equilibrium, with the incumbent sacrificing profits on the first buyer (who pays a price below the incumbent's marginal cost) and recouping losses on the second buyer (who will end up paying a much higher price). Predation will also be welfare-detrimental since efficiency would call for the rival firm to serve both new buyers.

Instead, if the entrant's efficiency advantage is large enough (that is, c_I is high), it will be the rival who makes the best price offer to the first buyer (since the incumbent offers the first buyer a price which is below its marginal cost, the rival's equilibrium price will also be strictly below c_I), and the second buyer will also buy from the rival (at the duopoly price, which will be just a shade below c_I).

The result that inefficient exclusion can arise in this setting, as well as in the more general model discussed below, can be interpreted as a result of the existence of *contracting externalities*. Contracting externalities exist when the payoff of a contracting agent depends not only on her own terms of trade, but also on the terms of trade obtained by other contracting agents.¹⁹ In this setting, the payoff of the incumbent depends not only on how much it sells to the first buyer, but also on how much the rival supplier sells to the first buyer which determines its second period cost and then the rents that the incumbent is able to extract from the second buyer. Then, exclusion of the more efficient rival takes place because, by removing (softening) competition in the second period, it allows the incumbent to extract larger revenues from the second buyer – that is, from the outsider of the first-period contracting – thereby increasing the joint payoff of the incumbent, the rival and the first buyer as compared to the situation where the rival supplies the first buyer. (On this, see also the discussion

¹⁹ See Segal (1999) for a general framework to study contracting externalities and to understand the sources of inefficiencies.

in Section 1.3.1.) Then, the incumbent is able to make an offer to the first buyer that the rival is unable to replicate, even though the rival is more cost-efficient over total production and submits bids at the same time as the incumbent. Further, since the agents contracting in the first period do not take into account the detrimental effect of their choice on the payoff of the outsider (that is, of the second buyer), the outcome that maximises their joint payoff turns out to be socially inefficient.²⁰

1.2.3.2 Generalisation of the example and extensions

The simple example described above is based on a number of simplifying assumptions. Hence, it is important to show that most of those assumptions are not crucial for the argument to work, but it is also important to identify the ingredients that are key for predation to be feasible and profitable. Otherwise, we would not be able to help formulate useful policy suggestions on how to assess allegations of predation.

We now discuss informally the robustness of the theory when changing some of the basic assumptions. (For a formal treatment, see the technical Section 1.3.)

The cost structure The key ingredient of the above example is that the incumbent is more efficient than the rival at supplying a *single* buyer (in supplying a single cohort of contestable buyers), while the rival is more efficient than the incumbent at supplying *both* (cohorts of) buyers.

As we hinted above, a cost structure of this type may arise if both the incumbent and the rival produce under technologies characterised by scale economies. The rival is endowed with a more advanced technology that allows it to produce at lower cost than the incumbent once achieved the efficient scale. The incumbent and the rival differ also in terms of captive buyers: on top of competing for the contestable buyers, the incumbent also serves a certain number of *captive* buyers, who bought from it in the past and are not willing to switch to another supplier (equivalently, one can imagine that the rival does not have a distribution network in some areas of the market). The rival, who may be a recent entrant, has also some captive buyers, but fewer than the incumbent. Alternatively, the smaller rival has not entered the market yet, and for this reason it has no captive buyers at all.

²⁰ In the terminology of Bernheim and Whinston (1998) this is a setting with ‘noncoincident market effects’. We further discuss the role of contracting externalities for inefficient exclusion in Chapters 2 and 3, when upstream suppliers compete for the first buyer offering conditional discounts or exclusive contracts. See also the deep discussion in Whinston (2006: Chapter 4.4).

Under scale economies, this *asymmetry in the number of captive buyers* may translate in the fact that serving a single new buyer is insufficient for the rival to reach efficient scale. Hence, the incumbent preserves an advantage over the rival when supplying a single contestable buyer; such an advantage is lost if the rival supplies both contestable customers and achieves efficient scale.

Another situation in which the above cost structure may arise is the one in which the *asymmetry* between the incumbent and the rival stems from *sunk investments*: the incumbent has already sunk an entry cost f when competition for the first contestable buyer starts, while the rival (an entrant firm in this case) has not. The entrant has a lower marginal cost than the incumbent, but it manages to cover the entry cost only if it supplies both buyers. Instead, the demand of a single buyer is insufficient to make entry profitable. There exist several real-world cases which resemble this situation. Indeed, there exist markets (such as public procurement markets) where some buyers may have to decide on the basis of tender offers, or where there may be large business customers which negotiate prices with their suppliers, *before* one or more suppliers have had the time (or ability) to develop the necessary production or sales capacity. Think, for instance, of a situation where the entry investment consists of building a large and complex infrastructure, carrying out construction work, or obtaining licenses and planning permissions. Also, there may be situations where the liberalisation process or government's auctioning of new technologies may entail sequential opening of market segments in a context in which some market participants are already incumbent whereas others are not. Indeed, in some of the cases that we discuss in Section 1.6, we shall see that the dominant firm's rival was a firm that had still to make investments to complete its network (telecom operators competing for public procurement and large buyers in *Telecom Italia – Comportamenti abusivi*), to build capacity for a new product (AMD, in the *Intel* case discussed in Chapter 2), to start new bus routes (2 Travel, in *Cardiff Bus*) or to establish a new readership and credibility with advertisers (Aberdeen Independent in *Aberdeen Journals*).

Finally, the above cost structure may also arise in environments characterised by learning-by-doing, where greater past production translates into lower current production costs through the accumulation of experience. One can think of a situation where the incumbent, who has been on the market for a longer period than the smaller rival, has accumulated *more learning*. This provides the incumbent with a cost advantage in supplying a single (cohort of) contestable buyer, whereas serving both (cohorts) of the

contestable buyers would allow the rival to fill the gap and produce more efficiently than the incumbent. Indeed, Cabral and Riordan (1994, 1997) study a duopoly model with endogenous learning-by-doing. They show that the incumbent may have an incentive to choose aggressive pricing in a first period to speed up learning, gain efficiency in the following period and, by stealing demand from the rival, to deny efficiency to the rival. If both effects are sufficiently strong, the rival may be induced to leave the market. However, differently from the model presented in this section, such an aggressive pricing policy is not necessarily welfare-detrimental, as the incumbent's acquired efficiency may benefit consumers and total welfare despite the exit of the rival. A recent paper, Besanko et al. (2014), isolates advantage-building and advantage-denying incentives for aggressive pricing in a fully dynamic model à la Cabral and Riordan (1994). They show, using numerical simulations, that the efficiency-denying incentive is the one leading to welfare-detrimental effects. In our model, the advantage-building motive is not indispensable, as the incumbent does not necessarily gain efficiency when it increases production during predation – see, for instance, the simple example discussed above in which the incumbent's marginal cost is constant. What is crucial in our model is the fact that predation denies efficient scale to the rival firm, thereby making it a weak competitor in the following period and generating rent extraction more favourable to the incumbent. And that is the precise reason why predation is welfare-detrimental.

An important implication of the above discussion is that the theory of predation presented in this section applies to situations in which the prey is a firm that is already in the market as well as to situations in which it is a new entrant.

Further, predation does not necessarily result in the prey exiting the market or being discouraged from entering the market. The prey may continue to operate and serve its captive buyers, but predation has allowed the dominant incumbent to marginalise its rival and relegate it to its niche market (that is, to its non-contestable buyers). The prey would have had the opportunity to expand its business and serve the new contestable buyers, but predatory pricing blocks this strategy.

These remarks are important for policy implications. First, because it is not necessary for the prey to exit the market in order for predation to be a feasible and profitable strategy for the incumbent. Second, because they show that the identification of contestable versus non-contestable (that is, captive) buyers may be crucial to appreciate the potential for exclusion in a given market: the larger the portion of contestable buyers

the more difficult for predation to be successful. Also, it can be shown (see Section 1.3.2.1) that the stronger the incumbency advantage – as captured by an increase in the number of the incumbent’s captive buyers – the more likely that predatory pricing will occur at equilibrium. This is because a larger non-contestable base makes the incumbent (weakly) more efficient in producing any of the two additional units. This, *ceteris paribus*, reduces the incumbent’s overall cost disadvantage and limits the amount of rents that the rival firm can extract from the second buyer if it takes the first one, thereby making it easier for the incumbent to win the competition for the first buyer. This has an important policy implication: abuse (here in the form of predation) is the more likely the stronger the dominant position of the incumbent. We shall come back to this point in Section 1.4.

Finally, let us note that predation may take place to the detriment of a rival which is already operating in the market also in situations where scale economies are on the demand side (that is, there are externalities among consumers), whether due to direct network effects or to two-sided markets effects, and incumbency advantages are due to a stronger customer base.

Strategic buyers In the basic version of the theory of predation based on scale economies and on an incumbency advantage, buyers (i) make independent decisions from one another and (ii) have to buy at exogenously given times. It is worth discussing what happens if these assumptions are relaxed.

(i) Joint decisions If buyers could take joint decisions, predation would not take place. Imagine, for instance, that buyers could delegate an agent to decide on the ground of their joint payoff. In such a case, the common agent would take into account that buying from the incumbent in the first period exerts a negative externality on the second period purchase by leading to higher second period prices. Using the terminology introduced by Bernheim and Whinston (1998), inefficient exclusion would not take place because all the agents would be represented in the first period negotiation and the negative externality exerted on the second buyer would be internalised. Alternatively, inefficient exclusion would not take place if buyers could pool their orders in a single period. For instance, the second buyer could ask the first buyer to purchase on her behalf as well, thereby purchasing two units in the first period. Now the first-period order would be sufficiently big to allow the rival to reach the scale it needs, and the most efficient rival would always end up getting this (large) order. A similar outcome would arise if the first buyer did not incur a loss in delaying

her purchase and both buyers could jointly decide in the second period.²¹ This means, for instance, that agreements among buyers to establish a central purchasing agency, or to delegate their purchase decisions to an entity which sets up a common public procurement system, or any other device which allows them to take joint decisions, are pro-competitive in the situation we describe. Such agreements can be interpreted as expression of *buyer power*, which would then limit the scope for predation.²²

(ii) Race to buy first (and simultaneous purchase decisions) Consider now the case where buyers do make independent decisions, but they are free to choose when to buy. Clearly, the first buyer would have no incentive to postpone her purchase because she obtains a higher surplus when buying first. However, the second buyer – if she could – would have an incentive to anticipate her purchase and be the first one to buy. Buyers will therefore engage in a race to be the first to buy. If there was an initial date before which purchases were not possible, both buyers would buy at that date.

It can be shown (see Section 1.3.2.2) that when the incumbent and the rival make simultaneous price offers to (independent) buyers exclusion could take place because of *buyer mis-coordination*. Consider again the simple example above, with the only variation that buyers receive simultaneous offers. In this case, it is possible that both buyers will end up buying from the incumbent even if it charges a higher price than its rival. Consider, for instance, a situation in which the incumbent sets the price f to both buyers 1 and 2, that is, $p_I^1 = p_I^2 = f$, while the rival sets a lower price to both of them: $p_R^1 = p_R^2 < f$. If one buyer expects the other buyer to choose the incumbent, then it has no incentive to deviate and turn to the rival, even if it offers a lower price. The buyer anticipates that its order alone is insufficient for the rival to achieve efficient scale so that the rival's cost to produce its unit alone (that is, f) exceeds the offered price. Then, the rival would prefer not to serve the deviant buyer.²³ In other words, if one buyer expects the other buyer to purchase from the incumbent, she would have no choice other than to purchase from the incumbent herself.

²¹ In both cases, though, the first buyer will want to be compensated by the second buyer, and will receive at least the same surplus as when decisions are decentralised, since she benefits from competition between suppliers when orders are made sequentially rather than at the same time.

²² Fumagalli and Motta (2008) arrive at the same conclusions. In that paper, though, as we discuss below (and analyse more formally in Section 1.3.2.2), buyers' (independent) decisions were simultaneous, and exclusion arose because of mis-coordination rather than because of the predatory mechanism highlighted here.

²³ Of course, mis-coordination problems do not arise if price offers represent an irreversible commitment to serve a customer.

Note that, in this context, the mechanism behind exclusion is due to buyers' inability to coordinate on the more efficient rival. For this reason, price below cost is not necessary for exclusion. In the case just discussed, both buyers buy from the incumbent, which is charging a price as high as f , simply because each of them expects the other to buy from the incumbent.²⁴ Note also that in this simultaneous-move game, for the same parameter values, there exist entry equilibria as well, where both buyers buy from the entrant (each buyer expects the other not to accept the incumbent's offer) and entry takes place. Unlike the sequential-purchase decisions, therefore, under simultaneous decisions exclusion may or may not occur for any given cost efficiency difference among the suppliers, and if exclusion does happen it is not clear what prices will emerge in equilibrium.²⁵

It is impossible to say *a priori* which environment among those discussed here would prevail in reality. Institutional features or legal constraints may explain the prevalence of a situation over another. Is it possible that buyers make joint decisions? Is it possible that buyers choose when to buy, and would they buy sequentially or simultaneously? For instance, legal constraints may prevent buyers from setting up joint purchases (maybe because regulators are afraid that if they agree on purchases they may also try to agree upon sales prices); the liberalisation process may be designed in such a way that a market would open before another; the existence of a patent may determine why a market may become contestable after another; some procurement rules may delay public procurement determining different purchase periods; financial constraints may delay purchase decisions of some consumers; and so on.

²⁴ Note that at this type of exclusionary equilibrium there are many prices which can be sustained by the incumbent. To be precise (technical remark): a continuum of prices can arise at equilibrium, each one supported by appropriate continuation equilibria concerning buyers' decisions.

²⁵ In cases where there is a multiplicity of equilibria, experimental evidence can provide some guidance for equilibrium selection. Experiments are studies, carried out in laboratories, where individuals are confronted with situations that seek to mimic the economic environment of interest. See Ochs (1995), for a survey on experiments on coordination games. The insights offered by Landeo and Spier (2009) and Boone et al. (2014) are particularly interesting for us. Even if in their framework the incumbent uses exclusive contracts to exclude, the nature of the coordination problem is the same as ours. These studies find that exclusion due to coordination failures occurs surprisingly often. Also, non-binding pre-play communication among buyers reduces the likelihood of exclusion. We discuss these findings more extensively in Chapter 3, where we consider exclusive contracts.

Buyer power The discussion so far has assumed that the two buyers (or markets) have equal size and are small, that is, their individual order is insufficient to guarantee efficient scale to the rival.

Let us discuss now the implications for the analysis of the *existence of large buyers*, whose individual demand is instead sufficient for the entrant to reach efficient scale. It is straightforward to show that predation cannot occur if both buyers are large. Similarly, there is no scope for predation if buyers are asymmetric and the second buyer is large: the incumbent will not be able to extract more rents than the rival from the second buyer, once secured the first one, and will have no incentive to bid more aggressively than the rival for the first buyer. Also from this perspective, then, *buyer power* limits the scope for predation.²⁶

A larger buyer in the second period can depict a market where the product is new and demand is expected to grow rapidly over time. This implies that in growing markets predation might be less of a concern. (For a more extensive analysis, see Section 1.3.2.3.)

Finally, the risk of predation is exacerbated if demand, in each of the two periods, is fragmented and buyers suffer from coordination failures. For simplicity, refer again to the simple example above. Imagine that in each period, instead of a single buyer with an order of size one, there are N buyers, each of them making an order of size $1/N$. In this context, as long as the rival needs more than $1 + 1/N$ orders to reach efficient scale, the incumbent may exclude the rival without charging below-cost pricing. The mechanism is similar to the one discussed for simultaneous offers: if a single first-period buyer expects the others to accept the incumbent's offer, then she has no incentive to address the rival – even if the latter offers a lower price – as its individual order added to the total second-period orders would not be enough for the rival to achieve efficient scale. Hence, buyer's fragmentation increases the likelihood of exclusion (by mis-coordination).

Downstream competition We have assumed so far that buyers are final consumers. This is not necessarily an innocent assumption in exclusionary models.²⁷

²⁶ Instead, predation is still possible if the large buyer comes first. If the rival is very inefficient in supplying the small buyer as compared to the incumbent, rent extraction in the second period would still be favourable to the incumbent. If this effect is strong enough, the incumbent can bid more aggressively than the rival for the first buyer and inefficient exclusion takes place.

²⁷ Indeed, Fumagalli and Motta (2006) – in the context of exclusive dealing – and Fumagalli and Motta (2008) – in the context of (simultaneous) price competition – show that

When buyers are downstream firms (for example, retailers) who are competing for consumers, we cannot assume any longer that the number of units they buy from their chosen supplier is fixed. In particular, in the extreme case where retailers are perceived as perfectly substitutable by final consumers, the buyer-retailer who pays the lower wholesale price will be able to win *all* the market demand (in our example, it will win both orders, not one). In turn, this means that even if the first buyer has committed to buy from the incumbent at a given wholesale price, the rival firm may guarantee itself enough scale to operate more efficiently than the incumbent by selling to the second buyer at a slightly lower price. Hence, even though the incumbent secured the first buyer, the rival does not suffer any disadvantage when competing for the second buyer, and the incumbent cannot benefit from more favourable rent extraction from the second buyer. In turn, this lack of advantage implies that the incumbent has no incentive to bid more aggressively than the rival for the first buyer. Note also that when competition is so fierce, the incumbent cannot recoup losses if it sells below cost to the first buyer. This buyer would dominate the downstream market and the incumbent could not make a profit on the second buyer. For these reasons, inefficient exclusion does not occur when there exists sufficiently fierce downstream competition.

If, instead, buyers-retailers were highly differentiated, or operated in different geographic markets (that is, downstream competition would be weak or absent), then exclusion might still occur: each retailer could bring only a share of the total market to the rival, and if the incumbent managed to win the first buyer, as long as the second buyer's order is not sufficient for the rival to reach efficient scale, the incumbent would act as a monopolist on the second buyer and would recoup the losses made on the first one.^{28,29}

Predation in markets with scope economies The mechanism described so far applies to several buyers belonging just as well to the same market as to *different* relevant (either product or geographic) markets, which are related by the existence of common costs or more generally by scope economies. Suppose that each of the two contestable buyers is a buyer (or

exclusion may not take place when buyers are retailers who compete fiercely enough in the downstream market.

²⁸ For a formal analysis, see Section 1.3.2.4.

²⁹ As we discuss in Chapter 3 downstream competition might have an ambiguous role in the context of exclusive dealing, and it is conceivable that in some circumstances it could actually facilitate exclusion. However, in this simpler setting where suppliers cannot offer exclusive contracts, fierce downstream competition would unambiguously prevent predation from taking place.

market) of a different product, with competition for buyer 2 taking place after competition for buyer 1 and with economies from joint production, for instance because of the presence of fixed costs common to the manufacturing of both products. In that case, the cost functions could be interpreted as total cost functions of the two products, and the interaction between scope economies and the incumbency advantage would lead to the rival being more efficient than the incumbent in supplying both products, but being less efficient when supplying the second product only. This latter assumption implies that the rival will be a weak competitor for the second buyer (or market) if the first is served by the incumbent in period 1. In turn, this determines that rent extraction in period 2, once the incumbent has secured the first buyer (or market), will be more favourable to the incumbent, and explains why the main results of our model carry over to this revised setting: the incumbent may act as a predator in the first market to preserve its dominant position in the other market.

Similarly, our mechanism applies and predation may arise in an environment where, in period 1, the rival can enter (or expand) in the market for the first product only, while in period 2 entry (or expansion) is possible in both product markets. The incumbent firm is already active in all of the markets. This may have been the case in some recently liberalised markets, such as postal services, where new entry is allowed in some segments of the market (mail-order parcel services and business-to-business mail), while the former public monopolist keeps a 'reserved area' for some period after liberalisation (for example, exclusive rights to carry letters and items weighing less than 200g); or it may be the case where tariffs or other barriers to trade are being phased out, or where it would take a long time to get all permits needed to operate locally, so that a new firm might be able to enter some markets immediately, but will be able to enter a particular foreign market only in the future.

The assumptions required in this case are that selling in both markets in period 2 is not enough for the rival to achieve efficient scale and become more efficient than the incumbent, while selling in the market for the first product in period 1 and in both markets in period 2 suffices. In this environment, it is easy to show that – in the presence of an incumbency advantage and economies of scope (in the postal service, a common distribution network that can be used to dispatch both letters and mail-order parcels, in the international markets example common R&D or technology) – the incumbent may predate in the markets which open first (for instance, the newly liberalised mail-order parcels' market), to preserve

its monopoly position in *all* the markets where it is active.³⁰ We will use this argument to discuss the actual *Deutsche Post* case in Section 1.6.6.

Network effects We have so far discussed our theory of predation in a context where scale economies are on the supply-side, but exactly the same mechanism applies to demand-side scale economies, due for instance to direct demand externalities (treated in this sub-section) or to two-sided markets effects (see the following sub-section).

Suppose that the incumbent and a rival firm produce two incompatible network products, with identical marginal costs. Each manufacturer has an installed base consisting of old customers who are not buying any longer, but who continue to use the (durable) network product they have bought in the past.³¹ The incumbent enjoys an *incumbency advantage* in that it can rely on a larger customer base than the rival. Suppose there are two *new* buyers who are considering buying one of the products, and whose utility increases with the number of (old and new) users of the network product they buy. This introduces a network externality which gives rise to (demand-side) scale economies: the more consumers a firm has, the more valued its network will be. Finally, suppose that the combination of network externalities and the incumbency advantage results in the following feature: even though at *full size* (that is, when both of the new buyers add to it) the quality of the rival's network is superior to the incumbent's, with only *one* new buyer the quality of the rival's product is inferior.

The reader will have noticed that this setting shares the same features as the general model described in Section 1.2.3.1, the only difference being that instead of assuming a relationship between number of units sold and cost-efficiency, we posit a relationship between number of units sold and perceived quality. One will therefore not be surprised that the game where the network firms compete sequentially for the two new buyers will produce a similar result to the examples described so far: if the quality gap between the rival's and the incumbent's network *at full size* is not too large, the incumbent will exclude the more efficient supplier by setting a price below its cost to the first buyer, and recouping the loss by charging the second buyer a much higher price. (Instead, for a sufficiently large quality

³⁰ This is very similar to the *defensive monopolisation* hypothesis which was first proposed by Carlton and Waldman (2002) in the context of a tying strategy inspired by the *Microsoft* case and of markets related by complementarity in consumption (rather than by the existence of common costs). We discuss this in detail in Chapter 4.

³¹ The same logic would apply to a case where the customer base is made of past customers who would buy again, but who have very high switching costs.

gap at full size, both buyers will buy from the rival, the first one of course at a much lower price than the second.)

The intuition behind this result is similar to the case of supply-side scale economies. Competition for the first buyer will be particularly intense because whoever secures the first buyer will supply also the second. The fact that at full size the quality of the rival's network is superior represents an advantage for the rival firm when competing for the first buyer. However the fact that one buyer is insufficient for the rival firm to reach a sufficient scale may allow the incumbent to extract more rents than the rival from the second buyer which – *ceteris paribus* – makes the incumbent more aggressive when competing for the first buyer. When this latter effect dominates, the incumbent secures the first buyer and excludes the more efficient rival.^{32,33} Similarly to the (general) model with supply-side scale economies, also in this case the stronger the incumbency advantage – that is, the more consolidated the customer base – the more likely predation arises in equilibrium. Again, this suggests that abuse is the more likely the higher the market share of the incumbent.

Despite its similarities, there are also some differences between the case of supply-side scale economies and the one with demand-side economies (due to the existence of network externalities between new and old consumers).

³² The paper by Farrell and Katz (2005) shares some similarities with our analysis. They also investigate price competition in an environment with network externalities. Both in our setting and theirs, denying sales to the rival in early periods weakens its ability to compete in later periods by making the rival's product less attractive to consumers. Below-cost pricing in early periods is then a natural outcome of price competition. However, the focus of the analysis is different in the two studies. Our purpose is to identify under which conditions below-cost pricing harms welfare by leading to the exclusion of a more efficient producer. Rather than attempting to separate 'predatory' from 'non-predatory' behaviour, their focus is instead on the effect of the imposition of price floors (such as a ban on below-cost pricing) on market outcomes and welfare. Their main finding is that whether such rules are welfare-detrimental or not depends on the way consumers form expectations on other consumers' behaviour. (Expectations do not play a role in our analysis because we assume that there is a single consumer in every period, which can be also interpreted as if there are many consumers who always manage to coordinate their actions.)

³³ Also in Carlton and Waldman (2002) – in the variant based on network externalities – the first cohort of consumers is the key one and competition for it may result in exclusion of the more efficient entrant. In their case, though, it is the fact that the incumbent is already active in the market for a complementary product to the network product that makes it more aggressive in bidding for the first cohort of customers. In turn, this occurs because the incumbent extracts the entire surplus generated by the system, if it dominates the market for the network product, while it is only partially able to do so if the entrant dominates such a market.

First of all, under network externalities the exclusion of the more efficient producer is not necessarily welfare-detrimental. The reason is that old customers, who are still using the incumbent's product, benefit when the new buyers join the incumbent's network. Their welfare gain may be large enough to dominate both the efficiency loss associated to the fact that new buyers use the inferior product and the loss suffered by the old customers of the rival due to the lack of expansion of their network. When this is the case, predatory pricing *excludes* the *more efficient* producer but is *welfare-beneficial*.³⁴ As the formal analysis of Section 1.3.2.7 will show, this situation is less likely to arise when the size of the incumbent's network is large enough to exhaust the externality generated by additional users, or to make it negligible. In this respect, the fact that the incumbent's dominance is pronounced, as proxied by the extent of its installed base of customers, makes welfare-detrimental predation more likely.

Two-sided markets Another instance of demand-side scale economies is given by two-sided markets. In such markets, a firm (or platform) typically sells its product or service to two different groups of consumers, each group (or side of the market) benefiting from positive externalities from the number of users on the *other* side. For example, credit card companies sell their services both to cardholders (shoppers who plan to use a card for their purchases) and to merchants (who accept cards as a way of payment). A cardholder's utility will typically increase with the number of merchants who accept her credit card (the card would have no value if no merchant accepted it), while a merchant's utility will increase with the number of users having a particular card. Other examples of two-sided markets are newspapers and yellow pages (who sell to readers and advertisers), game consoles (who sell to consumers and software developers), recruitment websites (matching firms and job-seekers), iPad and Kindle (competing for contracts with publishers on one side and for end-users on the other), pay-TV platforms (signing contracts with advertisers and content providers on one side, and with subscribers on the other), large music festivals (with fans valuing the quality and quantity of bands, and the latter being likelier to join provided the expected crowds are large enough or the venues are sufficiently prestigious), and so on.

In all these cases, a platform's success depends on its abilities to have both sides of the market 'on board'. As the economic literature well

³⁴ It is well known that in models with network externalities entry in the market by a new firm may be detrimental because it may lead to *stranding* (or reduced benefit from network externalities) of the old customers of the incumbent.

emphasises,³⁵ asymmetric price strategies are often followed: consumers on one side of the market may be enticed to join the platform at a very low price (or can indeed be subsidised), thereby making it attractive for the other side of the market to join that platform too.

In this case, an incumbent platform's advantage is given by the existence of a stronger installed customer base on both sides of the market. To the extent that platforms are incompatible, a rival firm may have a more attractive platform (or be more cost-efficient), but it will suffer from a (demand-side) scale disadvantage. Similarly to the network externality model sketched above, it is straightforward to construct examples where predation may arise because of the mechanism repeatedly mentioned above. (See Section 1.3.2.8 for a formal treatment.) A dominant platform may set very low prices to (or even subsidise) consumers on one side of the market in order to prevent the rival from achieving scale on that side, thereby also making it much less attractive for the other side to join the rival platform (and precisely because the rival platform is much less attractive, the incumbent may be able to set very high prices on the other side of the market, thereby recouping any losses made by preying on the first side).

It is worth noting therefore that – contrary to what is often suggested in the literature – low prices on one side are not necessarily an innocent strategy with pro-competitive effects. True, when a market is in its infancy, an asymmetric price strategy might be the key to ensure that the market will not fail (the two-sided externality may mean that nobody on one side buys, expecting nobody on the other side will – and zero or low prices on one side would break this self-fulfilling market failure); but when a market is already established, one might expect the market failure problem to be less important, and very low prices on one side might be an indication of an anti-competitive strategy aimed at excluding a rival platform.

As Section 1.3.2.8 will show formally, welfare considerations are very similar to the ones discussed in the case of network externalities.

In Section 1.6.1, we shall see how this framework can be used to rationalise predation in the well-known *Napp* case. There, *Napp* and its rivals were selling a pharmaceutical product to hospitals and to the 'community segment'. While hospitals' utility was not influenced by decision in the community segment, community decisions were heavily affected by hospitals', making this an (asymmetric) two-sided market.

³⁵ See Schmalensee (2002), Evans (2003), Rochet and Tirole (2003, 2006), Armstrong (2006a).

Relationships between traditional theories of predation and a theory based on scale economies and incumbency advantage The predation theory presented in this section neither invalidates nor generalises the traditional theories of predation – it is complementary to them. In some cases, the scale economies' mechanism we discuss might fit the evidence better. But in other cases, predation might be more likely motivated by the desire of an incumbent to build a reputation for aggressive behaviour or by the attempt of a well-funded dominant firm to make it more difficult for a new firm to obtain external funds. Further, these rationales might co-exist: the predation theory set out in this section does not rule out the possibility that an incumbent might want to deprive an actual entrant of the scale it needs while at the same time sending a message to other potential entrants that it is ready to do the same in the future; and being aggressive against an entrant might also have the effect of reducing the entrant's assets, and therefore making it more difficult for it to obtain funds in an imperfect capital market. (See the *ECS/AKZO* case discussed in Section 1.6.9.)

1.2.3.3 (Intertemporal) uniform prices and prohibition of below-cost prices

For predation to occur in a given relevant market, it must be that buyers will be charged different prices across periods, thus giving rise to some form of (*intertemporal*) *price discrimination*. If firms (or even only the dominant firm) were instead obliged to charge the same price in each period, then predation will never occur. Intuitively, the incumbent has an incentive to make losses on earlier buyers only if it can recoup them on later buyers, after it is clear that the prey will not be able to contest them. If (intertemporal) price discrimination were prohibited, this predatory strategy would not be possible: if the incumbent wanted to cut prices, it would have to do so for all buyers, thus implying that it would never want to sell below cost. Hence, the only equilibrium is such that the rival sets prices (slightly) below the marginal cost of the incumbent and it serves both buyers. (The proof can be found in Section 1.3.1.1.) At least in the case where scale economies are on the supply-side and the rival is more efficient than the incumbent over the entire contestable demand, it is straightforward to see that the same result would arise if a dominant incumbent was obliged never to sell below its marginal costs, as this would prevent it from undercutting the rival.

Leaving aside the practicability of these policies (for instance, supply and demand conditions change over time, so prohibiting price changes would

not be a sensible policy), they provide us with a natural benchmark for welfare analysis. How does this situation contrast with the outcome of the model we have discussed (where price discrimination or below-cost pricing are allowed)?

The answer – contrary to what one might expect at first sight – is ambiguous. To understand why, recall first of all that when the incumbent can price aggressively to some buyers, predation is not necessarily observed at equilibrium, depending on the efficiency gap between the incumbent and the rival.

Suppose the incumbent and the rival's production costs are sufficiently similar. In this case, a theory of predation based on scale economies and incumbency advantage predicts that predation will indeed take place. With respect to the *per se* rules such as banning price discrimination or below-cost pricing considered above, predation will be welfare-detrimental. This is because the more efficient rival is displaced by the less efficient one. Moreover, consumer surplus declines (the higher surplus enjoyed by the early buyer is outweighed by the lower surplus enjoyed by the later buyer). In our model with rigid demand functions such decrease in consumer surplus is perfectly compensated by the increase in the incumbent's profits. However, in a more general model with elastic demand, predation would decrease welfare also through a loss in consumer surplus.

Suppose instead that the rival is much more efficient than the incumbent. In this case, the entrant will be able to match the aggressive prices of the incumbent and will supply both new buyers at equilibrium, the first at a price below the marginal costs of the incumbent, and the second at a price equal to the incumbent's marginal cost. However, this implies that prohibiting price discrimination would not improve the price faced by the second buyer, but it would raise the price charged to the first buyer. In other words, the ban would chill competition and lead to (weakly) higher prices. In our model we adopt rigid demand functions, to simplify the exposition. Then, in this case where the efficiency gap between the entrant and the incumbent is large, a ban on price discrimination (or below-cost pricing) would reduce consumer surplus but would leave total welfare unchanged: entry occurs anyhow, and the loss in consumer surplus would be exactly compensated by higher producer surplus. In a more general setting with elastic demand functions, a ban would also produce an allocative inefficiency, thereby reducing total welfare.

Since it is impossible for a competition authority or a government to follow a policy contingent on the costs of the firms, the only conclusion we can arrive at is ambiguous. Measures aimed at discouraging price

aggressiveness by dominant firms would result in a trade-off. On the one hand, they would reduce the chances that anti-competitive exclusion would take place; on the other hand, when the entrant is sufficiently more efficient than the incumbent, exclusion would not occur and they would chill competition and result in higher prices.³⁶ We shall argue in Section 1.4 below that, if coupled with a plausible theory of harm consistent with the facts of the case, a dominant firm's below-cost pricing should be interpreted as strong evidence of anti-competitive conduct.

1.3 A Simple Theory of Predation*

In this section, we provide the analytical treatment of the theory of predation discussed in Section 1.2.3. We will start from a simple example (Section 1.3.1), and then we will prove that similar results arise in a more general setting (Section 1.3.2.1). Next, we will develop a number of extensions.

1.3.1 The Base Model (Supply-side Scale Economies)*

We assume that there are *two* buyers, B_1 and B_2 , with unit demand and valuation v for a homogeneous product.³⁷

An incumbent firm, I , and a rival firm, R , compete for the two buyers. The incumbent's marginal cost is constant and equal to $c_I > 0$. The rival's cost is f for the first unit and 0 for the second unit, with

$$c_I < f < 2c_I. \quad (1.1)$$

Assumption (1.1) ensures that two buyers are sufficient for the rival to achieve efficient scale and produce at lower costs than the incumbent, but a single buyer is not. We will discuss in the next section possible explanations for such property of the cost functions. Furthermore,

³⁶ See Karlinger and Motta (2012) for similar conclusions in a model with (simultaneous) price discrimination and network effects. See also Farrell and Katz (2005) for an analysis of the effects of the imposition of different forms of price floors in markets characterised by network externalities.

³⁷ The extension to n buyers would not create any conceptual difficulty and would leave qualitative results unchanged. By assuming elastic demands one would find similar qualitative results. The main difference would be that exclusion would entail not only a productive inefficiency but also an allocative inefficiency.

$f < 2c_I$ implies that it is socially efficient that the rival supplies the two buyers. We also assume that $f < v$.

Firms play the following game.

1. First period.

- (a) Firms I, R simultaneously set prices p_I^1 and p_R^1 to buyer 1.
- (b) Buyer 1 decides from whom to buy and the transaction takes place.

2. Second period.

- (a) Firms simultaneously set prices p_I^2 and p_R^2 to buyer 2.
- (b) Buyer 2 decides from whom to buy and the transaction takes place.

Proposition 1.1 (*Sequential – and discriminatory – offers*) *Equilibria of this game are as follows:*

- (**Exclusion**) *If $f > 3c_I/2 \equiv \tilde{f}_s$, then firm R and I set $p_I^{*1} = p_R^{*1} = f - c_I < c_I$, buyer 1 buys from I , firm R and I set the price $p_I^{*2} = p_R^{*2} = f$, the second buyer buys from I .*
- (**Entry/Expansion**) *If $f \leq \tilde{f}_s$, then firm R and I set $p_I^{*1} = p_R^{*1} = 2c_I - f < c_I$, buyer 1 buys from R , firm I and R set $p_I^{*2} = p_R^{*2} = c_I$ with the second buyer buying from R .*

Proof. Let us move by backward induction. Consider first the subgame following the first buyer choosing firm R . Then, in the second period, the rival's cost to supply B_2 is lower than the incumbent's: $0 < c_I$ by assumption. Standard Bertrand competition between cost-asymmetric firms takes place and the more efficient rival supplies the second buyer at the price $p_R^{*2} = c_I$ (here and in the rest of the book we disregard equilibria in weakly dominated strategies). If instead in the first period B_1 chose the incumbent, the rival's cost to supply B_2 is equal to f , while the incumbent's cost amounts to $c_I < f$ by assumption (1.1). Hence, in this case, it is the incumbent who is the low-cost supplier. In equilibrium the incumbent serves the second buyer at the price $p_I^{*2} = f$.

We now consider competition for the first buyer. Each firm anticipates that, by securing the first buyer, it will be able to supply the second buyer. Then, if it secures B_1 , the rival makes total profits $\pi_R = p_R^1 + c_I - f$.

This inequality identifies the minimum price at which firm R is willing to supply B_1 : $\tilde{p}_R^1 = f - c_I < c_I$ (by assumption (1.1)).

If instead B_1 decided to buy from the incumbent, then the incumbent's total profits are equal to $\pi_I = p_I^1 + f - 2c_I$. Then, the minimum price at which the incumbent is willing to supply B_1 is: $\tilde{p}_I^1 = 2c_I - f < c_I$ (by assumption (1.1)).

Differently stated, competition for the first buyer is like an asymmetric Bertrand case where the incumbent and the rival have, respectively, 'adjusted' costs $\tilde{c}_I = 2c_I - f$ and $\tilde{c}_R = f - c_I$, which correspond to the total cost of producing the two units (who supplies the first buyer, will supply also the second) minus the rents extracted from the second buyer. Note that the incumbent extracts more rents than the rival from the second buyer (i.e. $p_I^{*2} = f > c_I = p_R^{*2}$). Hence, even though the rival is more efficient than the incumbent in producing the two units (i.e. $f < 2c_I$), it is not necessarily the case that its adjusted cost is lower. Indeed, the incumbent's adjusted cost is lower if (and only if) $f > \frac{3c_I}{2} \equiv \tilde{f}_s$. The following situations can then arise:

(i) **(Exclusion)** If $f > \tilde{f}_s$ the equilibrium is such that $p_I^{*1} = p_R^{*1} = f - c_I$, and B_1 buys from the incumbent.

(ii) **(Entry/Expansion)** If $f \leq \tilde{f}_s$ the equilibrium is such that $p_I^{*1} = p_R^{*1} = 2c_I - f$, and B_1 buys from the rival. ■

Note that the exclusionary equilibrium arises even though the incumbent does not enjoy a first-mover advantage and the more efficient entrant can submit bids at the same time as the incumbent.³⁸ The source of exclusion is the interaction between scale economies and an incumbency advantage which leads to the incumbent being more efficient than the rival in producing a single contestable unit, even though the rival is more efficient than the incumbent in producing the two contestable units. It follows that, if the incumbent manages to serve the first buyer, the rival will be a weak competitor in the second period and the incumbent will charge a high price to the second buyer ($p_I^2 = f$). Instead, if the rival serves the first buyer, it will face tougher competition from the incumbent in the second period and it will be able to charge the lower price $p_R^2 = c_I < f$. This affects competition for the first buyer, where

³⁸ If the incumbent also enjoys a *first-mover advantage* exclusion will be easier. This is because the incumbent can take actions to attract the early buyer before the entrant can react, and can therefore exploit in the most profitable way the negative externality that the first buyer exerts on the other when it decides to buy from the incumbent.

firms anticipate that whoever supplies the first buyer will also supply the second. On the one hand, the fact that the rival is more efficient overall makes it more aggressive; on the other hand, the perspective of higher rent extraction makes the incumbent more aggressive. The previous proposition shows that if the (overall) efficiency advantage of the rival is not strong enough, then it is the incumbent which will make the winning bid for the first buyer. Therefore, predation will arise at the equilibrium and is welfare-detrimental.³⁹

Differently stated, exclusion of a more efficient rival allows the incumbent to extract larger rents from the second buyer (that is, from the outsider of the first-period contracting) thanks to a softening of competition in the second period. This maximises the joint payoff of the agents contracting in the first period (that is, the incumbent, the rival and the first buyer). This is the case both in the setting described here and in the more general model analysed below. Indeed, the condition that identifies when exclusion takes place can be interpreted along these lines. If the rival is excluded in period 1, the joint payoff of the contracting agents is given by:

$$\underbrace{v - p_I^{*1}}_{B_1} + \underbrace{p_I^{*1} + \underbrace{f}_{p_I^{*2}} - 2c_I}_{I} + \underbrace{0}_R \quad (1.2)$$

Instead, if the rival supplies the first buyer, the joint payoff of the contracting agents amounts to:

$$\underbrace{v - p_R^{*1}}_{B_1} + \underbrace{p_R^{*1} + \underbrace{c_I}_{p_R^{*2}} - f}_{R} + \underbrace{0}_I \quad (1.3)$$

³⁹ Another paper where exclusion may arise in the absence of a first mover advantage is Gans and King (2002). Differently from our setting, suppliers are perfectly symmetric and their focus is on asymmetries in contracting opportunities: there exist large buyers that can contract *ex ante* with suppliers and small buyers – whose demand is insufficient for a supplier to reach efficient scale – that can only trade *ex post* on a *single price* mass market. In this environment, it is in the interest of large buyers to commit *ex ante* to exclusivity with one supplier, to prevent the rival supplier from achieving the efficient scale. This will stifle competition in the mass market, thereby allowing to more rents to be extracted from small buyers. These rents are appropriated by large buyers through the *ex*

Comparing the two expressions, it turns out that the joint payoff of the three agents is larger under exclusion when $f > 3c_I/2$, which is precisely the condition central to Proposition 1.1.

When this condition is satisfied, the incumbent is able to make an offer to the first buyer that the rival is unable to replicate, even though the rival is more cost-efficient and submits bids at the same time as the incumbent. Further, since the agents contracting in the first period do not take into account the detrimental effect of their choice on the payoff of the outsider (that is, of the second buyer), the outcome that maximises their joint payoff turns out to be socially inefficient.

An entry deterrence interpretation One would obtain the same results in a model in which the rival is a new entrant that has not sunk the entry cost f when competition for the first buyer starts. The incumbent, instead, has already supplied past buyers and has already paid the entry cost. The two firms have constant marginal costs, with $c_R = 0 < c_I$. The timing of the game would be the same as the one described above, with the addition of an explicit entry decision for firm R at the end of each period⁴⁰ (and with the transaction with firm R taking place after the entry decision). We will fully develop the entry deterrence version of the model in Section 1.3.2.6.

1.3.1.1 (Intertemporal) uniform prices and prohibition of below cost prices*

The base model assumes that buyers can be charged different prices across periods, thus allowing for intertemporal price discrimination. The following Lemma shows that, if firms were instead obliged to charge the same price to each buyer, then predation would never occur. It is straightforward to see that the same result would arise if the incumbent was obliged to never set prices below marginal costs.

Lemma 1.2 (*Sequential – but uniform – offers*). *Under intertemporal uniform pricing, for all parameter values, firm R and I set $p_R^* = p_I^* = c_I$, both buyers buy from R .*

ante contracting. Allocative inefficiencies arise because small buyers pay too high a price, but there is no exclusionary intent in the suppliers' behaviour.

⁴⁰ Note, however, that if the rival enters at the end of the first period, it will not need to pay a fixed cost again. It is only if it does not enter in the first period, that it will have the chance to do it at the end of the second.

Proof. Since prices must be the same across periods, a firm cannot fix a price below cost in the first period recouping losses in the second period. The firms will therefore play the standard Bertrand game and $p_R^* = p_I^* = c_I$ is the equilibrium of the game, with the two buyers purchasing from the more efficient firm R . The incumbent has no incentive to undercut such a price because it should offer a price below cost to *both* buyers, which is unprofitable. In order to exclude the rival, the incumbent should offer a price $p_I < f/2$. Firm R would have no incentive to undercut such a price, but by assumption (1.1), $f/2 < c_I$ and the incumbent would make losses. ■

1.3.1.2 Welfare analysis*

The case of (intertemporal) uniform pricing provides us with the natural benchmark for welfare analysis.

Lemma 1.3 (i) *When $f > \tilde{f}_s$ and predation arises at the equilibrium consumers suffer as compared to the case where either (intertemporal) price discrimination or below-cost pricing is forbidden.*

(ii) *The predatory equilibrium is also welfare-inferior.*

(iii) *When $f \leq \tilde{f}_s$ and exclusion does not take place at the equilibrium, banning (intertemporal) price discrimination or below-cost pricing decreases consumer surplus.*

Proof. (i) If the incumbent is not allowed to price discriminate (intertemporally), at the unique equilibrium both buyers pay the price c_I . At the predatory equilibrium, buyers pay prices $p_I^1 = f - c_I$ and $p_I^2 = f$, respectively. The total price is lower in the former case, and thus total consumer surplus is higher, precisely when the condition under which predation takes place is satisfied: $2c_I < 2f - c_I$ when $f > 3c_I/2$. (ii) Since demands are rigid, total welfare at the predatory equilibrium is $2 - 2c_I$, while it is $2 - f$ when (intertemporal) price discrimination is banned. The assumption that firm R is more efficient than the incumbent over the two units (i.e. $f < 2c_I$) implies that the latter is larger. (iii) At the entry/expansion equilibrium, buyers pay prices $p_R^1 = 2c_I - f < c_I$ and $p_R^2 = c_I$, respectively. Since the first buyer pays a lower price as compared to the uniform price case, while the second buyer faces the same price, total consumer surplus decreases when (intertemporal) price discrimination is banned. Since demands are rigid and entry occurs anyhow, total welfare would be equal under price discrimination and under uniform pricing. ■

1.3.2 Generalisation of the Theory and Extensions*

In this section, we first present a more general setting where the interaction between supply-side scale economies and an incumbency advantage may give rise to predation (Section 1.3.2.1). We next also analyse in Sections 1.3.2.2-1.3.2.5 the robustness of our results to some of the simplifying assumptions we have adopted in the base model. Finally, we propose other settings where the same basic mechanism also applies: specifically, Section 1.3.2.6 analyses the case of products which develop over time, but are related because of common costs; Section 1.3.2.7 deals with scale economies on the demand side, created by the existence of network externalities; and Section 1.3.2.8 shows that predation may also appear in two-sided markets.

1.3.2.1 A more general model*

In the setting that we propose in this section, the incumbent's rival (denoted as R) may be interpreted either as a firm that is already in the market or as a new entrant but the asymmetry vis-à-vis the incumbent does not necessarily consist of the fact that the incumbent has already sunk the entry investment cost while the entrant has not.

We maintain the assumption that there are two contestable buyers/markets, B_1 and B_2 , each demanding one unit of an homogeneous good for any price (weakly) lower than v . We denote as $C_i(q_i)$ the total cost function of firm $i = I, R$, and we assume that firm R is *more efficient* than the incumbent in producing the *two* contestable units (assumption (1.4)), but is *less efficient* if it produces only one *unit* (assumption (1.5)):

$$C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) < C_I(\bar{q}_I + 2) - C_I(\bar{q}_I) \quad (1.4)$$

$$C_R(\bar{q}_R + 1) - C_R(\bar{q}_R) > C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) \quad (1.5)$$

where $\bar{q}_I > \bar{q}_R \geq 0$ denote the demand of some captive (that is, non-contestable) buyers/markets the two firms may possibly supply. Captive buyers may be past customers who have arbitrarily high switching costs and thus continue to buy from firm i , or buyers located in other geographical areas where firm i is active and which are separated by arbitrarily high transportation costs, or even past buyers whose choice affects present production costs, for instance due to learning-by-doing effects. Note that we assume that firm I benefits from an *incumbency*

advantage: it has been in the market for a longer period than the rival,⁴¹ or it has developed a more extended activity in other geographical areas, which translates in a *larger number* of captive buyers than the rival firm. Finally, we assume that $v > C_R(\bar{q}_R + 1) - C_R(\bar{q}_R)$, and that $C_R(\cdot)$ is strictly concave over the two contestable units, while $C_I(\cdot)$ is weakly concave.⁴²

The fact that the rival is less efficient than the incumbent on the first unit, in spite of being more efficient on the entire production, results from the interaction between the incumbency advantage discussed above and the existence of *scale/scope economies*. The fact that the incumbent supplies a higher number of captive customers may allow it to better exploit scale/scope economies and operate at lower incremental costs than the rival on the first contestable unit. Similarly, under learning-by-doing effects, an incumbent who has produced more in the past can produce an additional unit at lower costs.

Finally, we assume that the two buyers are approached *sequentially*, the timing of the game being as follows:

1. First period.

- (a) Firms I, R simultaneously set prices p_I^1 and p_R^1 to buyer B_1 .
- (b) B_1 decides from whom to buy and the transaction takes place.

2. Second period.

- (a) Firms simultaneously set prices p_I^2 and p_R^2 to buyer B_2 .
- (b) B_2 decides from whom to buy and the transaction takes place.⁴³

The subgame perfect Nash equilibria of this game are described by the following proposition:

Proposition 1.4 (*Sequential – and discriminatory – offers*) *There exists a threshold level C_P of firm R 's cost of producing the two units, with*

⁴¹ A natural interpretation is that the incumbent is the former monopolist in markets that have been recently liberalised.

⁴² Weak concavity of the incumbent's cost function simplifies the exposition. Indeed, we could allow $C_I(q_I)$ to be 'moderately' convex so as to ensure that a firm is more efficient in producing its second unit than the rival in producing its first unit. This property follows directly from assumptions (1.4) and (1.5) when the incumbent cost function is weakly concave.

⁴³ The results of the analysis would not change if both transactions took place at the end of the second period.

$C_P < C_I(\bar{q}_I + 2) - C_I(\bar{q}_I)$, such that:

- **(Predation)** If $C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) > C_P$, then the incumbent supplies both buyers. It sells below cost to the first buyer, while recouping losses on the second: $p_I^{*1} = \tilde{C}_R < C_I(\bar{q}_I + 1) - C_I(\bar{q}_I)$, $p_I^{*2} = C_R(\bar{q}_R + 1) - C_R(\bar{q}_R) > C_I(\bar{q}_I + 1) - C_I(\bar{q}_I)$.
- **(Entry/Expansion)** If $C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) \leq C_P$, then firm R supplies both buyers. The price paid by the first buyer is lower than the price paid by the second: $p_R^{*1} = \tilde{C}_I < C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) = p_R^{*2}$.

The threshold C_P is (weakly) decreasing in \bar{q}_I .

Proof. Let us move by backward induction. Let us consider first the subgame following B_1 choosing the incumbent. Standard Bertrand competition for the second buyer takes place, with the incumbent's cost to supply B_2 being lower than the rival's:

$$C_I(\bar{q}_I + 2) - C_I(\bar{q}_I + 1) \leq C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) < C_R(\bar{q}_R + 1) - C_R(\bar{q}_R), \quad (1.6)$$

the first inequality following from weak concavity of $C_I(\cdot)$ and the second from assumption (1.5). Hence, the incumbent serves the second buyer, at a price $p_I^{*2} = C_R(\bar{q}_R + 1) - C_R(\bar{q}_R)$. (Here, and in what follows, we disregard equilibria in weakly dominated strategies.)

Let us consider now the subgame following B_1 choosing the rival. In this case, the rival's cost to supply B_2 is lower than the incumbent's cost:

$$C_R(\bar{q}_R + 2) - C_R(\bar{q}_R + 1) < C_I(\bar{q}_I + 2) - C_I(\bar{q}_I + 1) \leq C_I(\bar{q}_I + 1) - C_I(\bar{q}_I), \quad (1.7)$$

the first inequality following from assumptions (1.4) and (1.5), the second from weak concavity of $C_I(\cdot)$. Hence, it is the rival that supplies the second buyer, at a price $p_R^{*2} = C_I(\bar{q}_I + 1) - C_I(\bar{q}_I)$.

Let us move to competition for the first buyer. Each firm anticipates that, by securing the first buyer, it will be able to supply also the second, thereby obtaining a total profit equal to:

$$\pi_i = p_i^1 + p_i^{*2} - (C_i(\bar{q}_i + 2) - C_i(\bar{q}_i)) \quad (1.8)$$

with $i = R, I$. We can thus denote as $\tilde{C}_i = C_i(\bar{q}_i + 2) - C_i(\bar{q}_i) - p_i^{*2}$, with $i = I, R$, each firm's adjusted cost to supply the first buyer, which corresponds to the total cost of producing the two units minus the rents extracted from the second buyer. Note that, by assumption (1.5), the

incumbent extracts more rents than the rival from the second buyer (i.e. $p_I^{*2} > p_R^{*2}$). Hence, even though the rival is more efficient than the incumbent in producing the two units, it is not necessarily the case that its adjusted cost is lower. More precisely, $\tilde{C}_R \leq \tilde{C}_I$ if and only if:

$$\begin{aligned} C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) &\leq C_I(\bar{q}_I + 2) - C_I(\bar{q}_I) - [C_R(\bar{q}_R + 1) \\ &\quad - C_R(\bar{q}_R) - (C_I(\bar{q}_I + 1) - C_I(\bar{q}_I))] \equiv C_P \end{aligned} \quad (1.9)$$

with $C_P < C_I(\bar{q}_I + 2) - C_I(\bar{q}_I)$ by assumption (1.5).

It follows that when $C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) > C_P$, the incumbent secures B_1 and sells at a price $p_I^{*1} = \tilde{C}_R$. If instead $C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) \leq C_P$, firm R secures B_1 and sells at a price $p_R^{*1} = \tilde{C}_I$.

Note that:

$$\begin{aligned} p_I^{*1} = \tilde{C}_R &= C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) - [C_I(\bar{q}_I + 1) - C_I(\bar{q}_I)] \\ &< C_I(\bar{q}_I + 2) - C_I(\bar{q}_I + 1) \leq C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) \end{aligned} \quad (1.10)$$

the first inequality following from assumption (1.4) and the second from weak concavity of $C_I(\cdot)$. Also:

$$\begin{aligned} p_R^{*1} = \tilde{C}_I &= C_I(\bar{q}_I + 2) - C_I(\bar{q}_I) - [C_R(\bar{q}_R + 1) - C_R(\bar{q}_R)] \\ &< C_I(\bar{q}_I + 2) - C_I(\bar{q}_I + 1) \leq C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) \end{aligned} \quad (1.11)$$

the first inequality following from assumption (1.5) and the second from weak concavity of $C_I(\cdot)$.

Weak concavity of $C_I(\cdot)$ also implies that the threshold C_P is weakly decreasing in \bar{q}_I . ■

Note that, from the last item of Proposition 1.4, the stronger the incumbency advantage – as captured by an increase in the number of the incumbent's captive buyers \bar{q}_I – the more likely the predatory equilibrium. This is because a larger \bar{q}_I makes the incumbent (weakly) more efficient in producing any of the two units. This, *ceteris paribus*, reduces the incumbent's overall cost disadvantage and limits the rival's rents extraction, thereby making it easier for the incumbent to win competition for B_1 .

Note also that the base model of Section 1.3.1 can be interpreted as a specific application of this general setting where:

$$C_R(\bar{q}_R + 1) - C_R(\bar{q}_R) = f \quad (1.12)$$

$$C_R(\bar{q}_R + 2) - C_R(\bar{q}_R + 1) = 0 \quad (1.13)$$

$$C_I(\bar{q}_I + 1) - C_I(\bar{q}_I) = C_I(\bar{q}_I + 2) - C_I(\bar{q}_I + 1) = c_I \quad (1.14)$$

Hence, assumptions (1.4) and (1.5) translate into:

$$c_I < f < 2c_I \quad (1.15)$$

Finally, the welfare considerations developed for the base-line model apply also to the general model.

1.3.2.2 Simultaneous (uniform) offers*

A crucial ingredient of the basic version of the model of predation based on scale economies and on incumbency advantage is that price offers to buyers are made sequentially. In this section, we analyse the case where offers to buyers are made simultaneously (and buyers choose simultaneously). We assume that prices are uniform across buyers, an assumption that we can rationalise and make consistent with the setting analysed so far by saying that only intertemporal discrimination is possible. (In Chapter 2, we consider the case of price discrimination within the same period.) We shall show that exclusion can still occur, but just because of *coordination failures among buyers*.⁴⁴

Let us rewrite the game as follows.

- a. Firms I, R offer uniform prices p_I, p_R to buyers B_1 and B_2 .
- b. Buyers independently and simultaneously decide from whom to buy (and are committed to their choice.)
- c. Firms decide whether to honour their order and transactions are made. (If a firm got orders from a buyer B_i at stage **b.**, but later it decides not to honour it, then stages **a.-c.** are repeated for B_i .)

Proposition 1.5 (*Simultaneous uniform pricing*) *The game admits two types of equilibria.*

- **Exclusionary (mis-coordination) equilibria.** *Firm I sets a price $p_I^* \in [c_I, f]$, firm R sets $p_R^* \leq p_I^*$, both buyers choose I.*
- **Entry/Expansion equilibria.** *Firm R sets $p_R^* \in [f/2, c_I]$, firm I sets $p_I^* \in [p_R^*, f]$, both buyers choose R.*

⁴⁴ As we discuss in Chapter 3, coordination failures are the source of inefficient exclusion also in a model where buyers are offered exclusive contracts. See Rasmusen et al. (1991) and Segal and Whinston (2000).

Proof. The proof is by backward induction. First of all, note that at stage c, if firm R has received an order and decides not to honour it, then the incumbent would set $p_I = f$ and serve the buyer. Note also that at stage c firms decide to honour orders only if they have collected enough revenues to cover production costs.

Let us now consider the buyers' game at stage b, given the prices bid at stage a. To understand the logic of the proof it suffices to focus on the following price configurations:⁴⁵

- If $p_I < p_R$, it is easy to see that there is a unique equilibrium where both buyers choose the incumbent firm.
- If $f/2 \leq p_R < p_I \leq f$, there are two equilibria in the buyers' game: the first one, where both buyers choose firm R and the second where both buyers choose firm I . Let us consider first the equilibrium where both buyers choose firm R . Since $2p_R \geq f$, firm R will honour its orders, and B_i will pay the price p_R . She has no incentive to deviate and choose the incumbent, as she would pay the higher price p_I . Let us consider the second equilibrium, where both buyers choose firm I . Given that B_j buys from I , B_i has no incentive to deviate and choose R . If she did so, firm R would not honour the order as her demand alone is insufficient to make firm R cover production costs ($p_R < f$), and she would then be obliged to buy from the incumbent at the price $f \geq p_I$. (The argument is the same when $p_R = p_I < f$. When $p_R = p_I = f$ any buyers' choice is an equilibrium.)
- If instead $p_R < p_I$ with $p_I > f$ the unique continuation equilibrium is such that both buyers choose firm R . Now choosing the rival is a dominant strategy for any buyer: she will pay a lower price both if the rival honours the order (since $p_R < p_I$) and if the rival does not and she will buy the good later from the incumbent ($f < p_I$).

We can now move to stage a. We characterise the equilibrium solutions. According to the continuation equilibria following the bids where $p_R \leq p_I \leq f$ exclusion may either occur or not.

Consider first the **mis-coordination equilibria**. The pair of prices $p_I^* = f$ and $p_R^* \leq f$ is sustained as an equilibrium by having both buyers choosing the incumbent following any bid where $p_R \leq p_I = f$. Such

⁴⁵ For a complete proof see Fumagalli and Motta (2008). Even though it analyses a setting with n buyers and elastic demand, the logic of the proof is the same.

continuation equilibria ensure that firm R has no incentive to deviate and decrease its price, as this would not attract buyers. Clearly, firm R has no incentive to increase its price as the unique continuation equilibrium is such that both buyers would choose the incumbent. Turning to firm I , it has no incentive to decrease its price. It has no incentive to increase its price either as it would lose both buyers.

Mis-coordination equilibria also exist where $p_I^* = p < f$ and $p_R^* \leq p$. They are sustained by having both buyers choosing the incumbent following any bid where $p_R \leq p_I = p$, while both buyers choosing the rival following any bid $p_I > p$ and $p_R \leq p_I$. These continuation equilibria ensure that the incumbent has no incentive to deviate and bid a price above p because it would lose all buyers; also, the rival has no incentive to change its bid because this would not allow to attract buyers nor to make sales profitable.

Finally, a mis-coordination equilibrium where $p_I > f$ does not exist. Firm R would have an incentive to deviate and slightly undercut the incumbent as this allows it to capture both buyers.

Let us turn now to **entry/expansion equilibria**. First, firm R cannot supply the buyers at the equilibrium if it bids a price $p_R > c_I$: the incumbent could profitably undercut and obtain all buyers. Firm R cannot supply the buyers at the equilibrium if it bids a price $p_R < f/2$ either: the revenues from both buyers are not enough to cover the production costs.

Equilibria where $p_R^* = p \in [f/2, c_I]$ and $p_I^* = p$ are sustained by having both buyers choosing the rival following any bid where $p_R < p_I$. The rival cannot deviate by increasing its price as it would lose all orders. In turn, the incumbent is indifferent between p and any higher price because no buyer would patronise it in any case; instead, it captures both buyers by decreasing its price but it would not break even as the deviation price would be below c_I .

Finally, there are also entry/expansion equilibria where $p_I > p_R$: $p_R^* = p \in [f/2, c_I]$ and $p_I^* \in (p, f]$. They are sustained by having both buyers choosing the rival following any bid where $p_I > p_R = p$ and both buyers choosing the incumbent following any bid where $p < p_R \leq p_I$. The latter ensures that firm R cannot increase its payoff by increasing the price and setting it equal or lower than the incumbent's because it would lose all the buyers. ■

1.3.2.3 Growing markets*

In our base model, we assume that buyers (markets) have the same demand (or equivalently that markets are equally sized). To understand how the relative importance of buyers affects the results, consider the same setting as in the base model, but assume that buyer B_1 's demand is $1 - k$, and B_2 's demand is $1 + k$, with $0 \leq k < 1$, so that the total market size does not change and is still equal to 2. The base model is embedded here for $k = 0$. We assume in this case that the incumbent has a constant marginal cost c_I , while the rival bears a total cost f on the first $1 + k$ units and a total cost equal to 0 on the subsequent $1 - k$ units, with:

$$(1 + k)c_I < f < 2c_I. \quad (1.16)$$

Note that the condition that ensures that the demand of the second buyer is insufficient for the rival to achieve efficient scale becomes more stringent, since the demand of that buyer has increased. Also, as the following Proposition shows, as the second buyer becomes more important, exclusion is less likely to take place.

Proposition 1.6 *If the second buyer is more important relative to the first one, exclusion of the more efficient rival arises if (and only if) $f > \tilde{f}(k) \equiv (3 + k)c_I/2$. The threshold $\tilde{f}(k)$ is increasing in k .*

Proof. Let us move by backward induction. If B_1 decided to buy from the incumbent at the price p_I^1 , then the incumbent wins the second buyer as it is more efficient than the rival in supplying B_2 : $f > c_I(1 + k)$ by assumption. This means that, if the incumbent secures the first buyer offering the price p_I^1 , then it makes total profits equal to $\pi_I(p_I^1) = p_I^1(1 - k) + f - 2c_I$. The incumbent's minimum price to supply the first buyer is $\tilde{p}_I^1 = (2c_I - f)/(1 - k)$.

If instead B_1 decided to buy from firm R , the rival is more efficient than the incumbent in supplying the second buyer. Indeed, the total cost of the rival in supplying B_2 , once it has already supplied B_1 , is $2kf/(1 + k)$ which is always lower than the incumbent's cost to supply B_2 :

$$\frac{2kf}{1 + k} < (1 + k)c_I \Leftrightarrow f < \frac{(1 + k)^2 c_I}{2k} \quad (1.17)$$

which is always satisfied since $(1 + k)^2/(2k) \geq 2$ for $k \in [0, 1]$ and $f < 2c_I$ by assumption (1.16). Then, if B_1 decided to buy from the rival at the price p_R^1 , the total profits that firm R makes are: $\pi_R(p_R^1) = p_R^1(1 - k)$

$+(1+k)c_I - f$. The rival's minimum price to supply the first buyer is $\tilde{p}_R^1 = f - (1+k)c_I/(1-k)$.

Combining the above results one obtains that $\tilde{p}_I^1 < \tilde{p}_R^1$ if (and only if) $f > \frac{(3+k)c_I}{2}$. ■

In the simple example just described, the rents that the incumbent extracts from the second buyer are pinned down by f and do not vary with the second buyers' size, whereas the rents that the rival extracts from B_2 do increase as B_2 's size expands. This leads to the unambiguous conclusion that predation becomes less likely as the second buyer becomes more important.

In a more general setting, the implications are less clear-cut. To see why, recall that a necessary condition for (inefficient) exclusion is that the $1+k$ units are insufficient for the rival to reach the efficient scale and produce more efficiently than the incumbent:

$$C_R(\bar{q}_R + 1 + k) - C_R(\bar{q}_R) > C_I(\bar{q}_I + 1 + k) - C_I(\bar{q}_I), \quad (1.18)$$

that is what allows the incumbent to extract more rents than firm R from the second buyer, once the first one has been secured, which in turn is necessary for the incumbent to have an incentive to bid more aggressively for B_1 . When $k = 1$, the above condition cannot be satisfied as it would contradict assumption (1.4), which ensures that firm R is more efficient than the incumbent on the entire production and thus that exclusion (if any) is welfare-detrimental. Instead, by assumption (1.5), the above condition is satisfied when $k = 0$ and buyers are symmetric. By continuity, there exists a critical size of the second buyer $1+k^*$ such that the above condition does not hold and thus inefficient exclusion cannot arise if the size of the second buyer is above the threshold level.

Now, when condition (1.18) is satisfied, following the same logic of Section 1.3.2.1, one can easily show that predatory pricing and inefficient exclusion take place if (and only if) firm R 's cost advantage is not too large, that is, iff $C_R(\bar{q}_R + 2) - C_R(\bar{q}_R) > C_P(k)$ where

$$C_P(k) \equiv C_I(\bar{q}_I + 2) - C_I(\bar{q}_I) - [C_R(\bar{q}_R + 1 + k) - C_R(\bar{q}_R) - (C_I(\bar{q}_I + 1 + k) - C_I(\bar{q}_I))]. \quad (1.19)$$

Note that, without imposing specific restrictions on the slope of the cost functions, one cannot tell whether inefficient exclusion becomes more or less likely as buyers' asymmetry increases, that is, as k increases. Indeed,

an expansion of the second buyer's demand allows both suppliers to extract more rents from B_2 , once secured B_1 , thereby inducing a more aggressive bidding for the first buyer by both suppliers. The only possible claim is that for values of k sufficiently close to k^* the threshold $C_P(k)$ is increasing in k , and thus exclusion becomes less likely as the second period demand expands.

1.3.2.4 Downstream competition*

To formalise the situation where buyers compete downstream, keep the same assumptions on parameters as in the base model, but assume that in each market $i = 1, 2$ there is a mass of consumers normalised to 1 and with unit valuation and demand for a certain product. Consumers cannot be served directly by firms I and R , but only by retailers. Consider the following timing of the game:

1. In the first period, firms I, R set wholesale prices w_I^1 and w_R^1 to retailer B_1 , who decides from whom to buy (but does not commit on the size of the order).
2. In the second period, firms simultaneously set prices w_I^2 and w_R^2 to retailer B_2 , who decides from whom to buy.
3. In the third period, retailers set prices p_1 and p_2 . Consumers in each market decide. Transactions take place.

In what follows, we limit ourselves to state the result for two extreme cases: (a) independent markets: consumers in market i can buy only from retailer B_i ; (b) perfect substitutes with Bertrand competition: consumers can buy from either retailer or both.⁴⁶ For intermediate competition cases, we would expect that – as in Fumagalli and Motta (2008) – if there is sufficiently fierce competition downstream, predation will not take place in equilibrium.

Proposition 1.7 (*Downstream competition*) *Suppose buyers are retailers who sell to final consumers. Equilibria of this game are as follows:*

- (**Independent market areas**) *If each retailer sells in a separate final market of size 1 and with unit valuation, then the equilibria are the*

⁴⁶ One can rationalise the two cases as due to transportation costs. If retailer B_i is located in market i , the independent markets case corresponds to segmented geographical markets

same as in the base model (in particular, predation arises if $f > 3c_I/2 \equiv \tilde{f}_s$).

- (Fierce competition) If the two markets are integrated and retailers compete in prices for final customers, exclusionary equilibria do not exist. Entry/Expansion equilibria exist and are such that firm R supplies one or both retailers at a wholesale price $w_R = c_I$.

Proof. (Independent market areas) If retailers are selling in independent markets, then each retailer can sell at most one unit of the product. Hence, everything will be as in the base model where buyers are final consumers who buy at most one unit.

(Fierce competition) Let us move backwards. In the third period, standard Bertrand competition between retailers takes place. Note that, for any w^1 and w^2 paid to upstream suppliers, the low-cost retailer (if any) captures the entire downstream market and sells two units of the product. In the second period, for given w^1 set for the first retailer, the incumbent and the rival compete for the second retailer. Let us distinguish the following cases:

1. Let us consider first the case where $w^1 > c_I$. Then, irrespective of whether the first buyer committed to buy from the incumbent or the rival, competition for B_2 will result in the second retailer buying from firm R at a price $w_R^2 = c_I$ (or slightly below) and selling *two* units to final consumers. Firm R covers its total production costs as $f < 2c_I$.
2. Let us consider now the case where $w^1 \in (f/2, c_I]$. If the first retailer committed to I , then in the second period the second retailer will pay a price w_R^2 slightly below w^1 to firm R and will dominate the final market selling *two* units to final consumers. If the first retailer committed to R , then any $w_I^2 \geq w^1$ and $w_R^2 \geq w_I^2$ is an equilibrium, with firm R either selling two units to the first retailer or one unit to each retailer (when $w_R^2 = w_I^2$). In all the cases it is firm R that sells the input and cover its total production costs.
3. Finally, let us consider the case where $w^1 \in [0, f/2]$. If the first buyer committed to buy from firm i (with $i = I, R$), then firm j (with $j = I, R \neq i$) has no incentive to secure the second buyer. In order to sell a positive amount of input, firm j should offer a price $w^2 < w^1$

with arbitrarily large transportation costs, the perfect competition case corresponds to integrated markets with zero transportation costs.

to the second buyer, which would allow the latter to dominate the downstream market and sell two units to final consumers. However, firm j would not cover its production costs, as $w^2 < w^1 < f/2 < c_I$. Then, any $w_i^2 \geq w^1$ and $w_j^2 \geq w^1$ is an equilibrium with firm i either selling the two units to the first retailer or one unit to each retailer (when $w_i^2 = w^1$).

Let us now consider price offers to the first buyer. Firms anticipate that for any $w^1 > f/2$, firm R will end up selling the two units, with the final market being dominated by either of the retailers, or being shared by both of them. In order to exclude the rival, the incumbent should capture the first buyer by offering $w_I^1 \leq f/2 < c_I$. However, the incumbent makes losses if it offers such a price to the first retailer: even though the rival is excluded, the second retailer will not sell unless it pays a price $w^2 \leq w^1$; hence the two units would be sold below costs either to the first retailer (if I chooses $w_I^2 > w_I^1$) or to both (if I chooses $w_I^2 = w_I^1$).

Hence, an equilibrium where the rival is excluded from the market cannot exist. It is easy to see that equilibria involve many different combinations of wholesale prices, with the incumbent making zero profits in each of them and the entrant selling the two units at a wholesale price which cannot exceed $w_R = c_I$.

Note the role played by fierce downstream competition. Even if the first buyer has committed to buy from the incumbent at a certain wholesale price, the rival firm can guarantee itself enough scale to operate more efficiently than the incumbent by selling to the second buyer at a lower price. Hence, when downstream competition is fierce, even though the incumbent secured the first buyer, firm R does not suffer any disadvantage when competing for B_2 and the incumbent cannot take advantage of more favourable rent extraction from the second buyer. Moreover, when competition is so fierce, the incumbent cannot recoup losses if it sells below cost to the first buyer: this buyer would dominate the downstream market and the incumbent could not make a profit on the second buyer. For these reasons, inefficient exclusion cannot arise at the equilibrium. ■

1.3.2.5 Renegotiation (or breach of orders)*

The base model of Section 1.3.1 assumes that transactions take place in each period, immediately after the buyer has chosen the supplier. Imagine, instead, that in each period buyers decide from whom to buy

(with their decision having commitment value), whereas transactions take place at the end of the two periods. The predatory equilibrium in which both buyers choose the incumbent even though the rival could supply the two units at lower costs may arise also in this context. However, the fact that transactions take place at the end raises the question of whether the predatory equilibrium would survive to the possibility of renegotiating the buyers' decisions. There exist contexts where there might be little scope for renegotiation. For instance, renegotiation might require some form of agreement/coordination between suppliers and antitrust laws might prohibit or impose restrictions on this type of behaviour. Alternatively, renegotiation costs might be high because breaching the initial decision may involve substantial legal costs or because of the costs of delaying consumption and production until a new agreement is reached. In an environment where, instead, renegotiation costs are sufficiently low, an equilibrium where both buyers choose the incumbent might still arise – sustained by the incumbent's ability to extract part of the gain from renegotiation – but it would not involve exclusion of the more efficient supplier: it is the rival that ultimately serves the customers, following a payment from the customers to the incumbent in the renegotiation phase.

1.3.2.6 Predation in markets with common costs*

In this section, we present a slightly modified version of the base model of Section 1.3.1. First, we will assume that the rival is an entrant firm that has not paid the entry cost yet when competition in the first period takes place. Instead the incumbent is an established firm that has already sunk the entry cost. The analysis will show that the mechanisms leading to exclusion are the same as in the base model. Second, we assume here that there are two *distinct* markets denoted as L and M . There is independence on the demand side. There is one consumer in market L and one in market M , and each of them attaches a unit valuation to the product. Third, the evolution of the markets is the following. Market M is the market which opens first to competition (for instance, in the postal service, market L is the market for letters, market M that for mail-order parcels): in the first period firm E can enter only market M ; in the second period firm E can enter both market M and market L . This timing reflects the situation created by liberalisation processes which have opened to competition some segments of a given market

before others. The discussion of the *Deutsche Post* case in Section 1.6.6 refers to this version of the model.

The entry cost f is common to the two markets: once paid f for entering market M , a firm does not need any other set-up costs. In both markets, marginal costs are given by $c_E = 0 < c_I < 1/2$. Finally, we assume that

$$2c_I < f < 3c_I \quad (1.20)$$

The game is as follows:

1. First period.

(a) Firms I, E simultaneously set prices $p_I^{L,1}$, $p_I^{M,1}$ and $p_E^{M,1}$ to the buyers, who decide if and from whom to buy. (b) Firm E decides whether to enter market M (and pay f) or not. (c) Transactions take place. If E got the order in market M but did not enter, the buyer purchases from I at the offered price $p_I^{M,1}$.⁴⁷

2. Second period.

(a) Firms I, E simultaneously set prices $p_I^{L,2}$, $p_I^{M,2}$ and $p_E^{L,2}$, $p_E^{M,2}$ to the buyers, who decide if and from whom to buy. (b) Firm E decides whether to enter in either market L , or market M , or both. If it has not entered market M yet, by paying the cost f firm E can enter both markets. If it has already entered market M , it does not need to pay any additional set-up cost to operate in market L . (c) Transactions take place. If E got an order in a market but did not enter that market, the buyer purchases from I at the offered price $p_I^{j,2}$ with $j = M, L$.⁴⁸

In what follows we show that if the fixed cost f is large enough there will be a predatory equilibrium with deterred entry; otherwise, firm E will enter market M in the first period, and market L in the second. As in the base model, there is scope for predation because the incumbent – but not firm E – has already sunk the common fixed cost when offers are

⁴⁷ The results would not change if we assumed that the buyer whose order remains unfulfilled is forced to buy from the incumbent which would then charge the monopoly price.

⁴⁸ Allowing the entrant to enter also at the end of the second period only affects the ‘maximum’ price that firm I could charge in the second period when firm E did not enter in period 1. Allowing for a second chance of entry implies that the incumbent will charge the limit price $p_I^{L,2} + p_I^{M,2} = f$ (if the price was higher, the entrant would undercut the incumbent and would enter) instead of the monopoly price $1 > f$ in each market. Hence,

made, and because demand in the second period alone is insufficient for firm E to reach efficient scale.

Proposition 1.8 (*Predation in sequential markets with common costs.*)
Equilibria of this game are as follows:

- **(Predation)** If $f > 5c_I/2$ then: in the first period, firm E and I sets $p_E^{M,1} = p_I^{M,1} = f - 2c_I < c_I$, $p_I^{L,1} = 1$, firm I supplies both markets and E does not enter M . In the second period, firm E and I set $p_E^{L,2} = p_I^{L,2} = z$, $p_E^{M,2} = p_I^{M,2} = f - z$ (with $z \in [0, f]$), and I supplies both markets. Firm E does not enter any market.
- **(Entry)** If $f \leq 5c_I/2$ then: in the first period, firm E and I set $p_E^{M,1} = p_I^{M,1} = 3c_I - f < c_I$, I sets $p_I^{L,1} = 1$, E supplies market M , I market L . In the second period, firm E and I set $p_E^{L,2} = p_I^{L,2} = p_E^{M,2} = p_I^{M,2} = c_I$, and E supplies both markets.

Proof. By backward induction. Consider the second period first. If E already entered market M in period 1, then it does not have to incur any cost to enter market L . Standard Bertrand competition with cost-asymmetric firms takes place, and the more efficient firm E supplies both markets fixing the price $p_E^{L,2} = p_E^{M,2} = c_I$. If E did not enter M in period 1, then it has still to pay the common entry cost when it competes with the incumbent in the second period. Since $f > 2c_I$ (by assumption (1.20)), in equilibrium firm I sets any pair of prices that satisfies $p_I^{L,2} + p_I^{M,2} = f$ and entry will not occur.

Consider now the first period. Since second-period rents alone are insufficient to cover the entry cost, if the incumbent captured market M in the first period, firm E does not enter. Hence, the incumbent will dominate both markets in the second period (at a total price f), thereby making total profits $p_I^{M,1} + 1 - 2c_I + f - 2c_I$. The incumbent's minimum price to supply market M in the first period is thus $\tilde{p}_I^{M,1} = 3c_I - f < c_I$. (Note that the incumbent supplies market L in the first period at the monopoly price $v = 1$ irrespective of whether it secures market M .) If firm E captures market M in the first period, then it finds it profitable to pay f if the rents collected in market M are large enough: $p_E^{M,1} + 2c_I - f \geq 0$.

by assuming that entry is possible also in the second period we make it more difficult for exclusion to take place.

The minimum price that firm E is willing to offer to supply market M in the first period is then: $\widetilde{p}_E^{M,1} = f - 2c_I < c_I$ (by assumption (1.20)). It is easy to check that when $f > 5c_I/2$ it is the incumbent who is willing to bid more aggressively for market M in the first period. ■

1.3.2.7 Network effects*

In this section, we assume that the incumbent and the rival are equally efficient in producing two differentiated and incompatible network products, and have a constant unit cost equal to c . Each manufacturer has an installed base of customers b_i with $i = I, R$, that is, old customers who are not buying any longer, but continue to use the network product. Also in this case, we assume that the incumbent enjoys an *incumbency advantage* and can rely on a larger installed customer base than the rival: $b_I > b_R \geq 0$. There are two *new* buyers, B_1 and B_2 , who enjoy utility $U_i = v_i(n_i) - p_i$ if they buy one unit of the network product from firm $i = I, R$, where $n_i \in N^+$ indicates the total number of users (including present and past buyers). There are direct network externalities in that the utility enjoyed by a user of network i increases with the total number of users of that network: $v_i'(n_i) \geq 0$. Even if not necessary for our results, we also assume that $v_i''(n_i) \leq 0$. Finally, similarly to the analysis of Sections 1.3.1 and 1.3.2.1, we assume that the combination of network externalities and the incumbency advantage results in the following feature: even though at *full size* (that is, when both of the new buyers add to it) the quality of the rival's network is superior to the incumbent's (assumption (1.21)), with only *one* new buyer the quality of firm R 's product is inferior (assumption (1.22)):

$$v_R(b_R + 2) > v_I(b_I + 2) \quad (1.21)$$

$$v_I(b_I + 1) > v_R(b_R + 1) \quad (1.22)$$

The game is as follows.

1. First period.
 - (a) Firms I, R simultaneously set prices p_I^1 and p_R^1 to the first buyer.
 - (b) B_1 decides from whom to buy.
2. Second period.
 - (a) Firms I, R simultaneously set prices p_I^2 and p_R^2 to the second buyer.
 - (b) B_2 decides from whom to buy.

3. Third period.

Consumption takes place and utilities are realised.

The following Proposition shows that also in this case – if the quality gap between the rival's and the incumbent's network at full size is not too large – by pricing below cost, the incumbent can exclude the more efficient supplier.

Proposition 1.9 *There exists a threshold level v_P of the utility of firm R's network, with $v_P > v_I(b_I + 2)$ such that:*

- **(Predation)** *If $v_R(b_R + 2) < v_P$, then the incumbent supplies both buyers. It sells below cost to the first buyer, while recouping on the second buyer: $p_I^{*1} = \tilde{c}_R - [v_R(b_R + 2) - v_I(b_I + 2)] < c$ and $p_I^{*2} = c + v_I(b_I + 2) - v_R(b_R + 1) > c$.*
- **(Entry/Expansion)** *If $v_R(b_R + 2) \geq v_P$, then firm R supplies both buyers. The price paid by the first buyer is lower than the price paid by the second: $p_R^{*1} = \tilde{c}_I + [v_R(b_R + 2) - v_I(b_I + 2)] < c + v_R(b_R + 2) - v_I(b_I + 1) = p_R^{*2}$.*

The threshold v_P is (weakly) increasing in b_I .

Proof. Let us move by backward induction. The outcome of competition for the second buyer, B_2 , depends on the choice made by the first one. Let us consider first the subgame following B_1 choosing the incumbent. From assumption (1.22) and from $v_I(n_i)$ being (weakly) increasing in the total number of users, it follows that the quality of the incumbent's network when B_2 joins is superior to the quality of the rival's network when B_2 joins:

$$v_I(b_I + 2) \geq v_I(b_I + 1) > v_R(b_R + 1) \quad (1.23)$$

Hence, in order to attract B_2 , the rival should discount the incumbent's price by an amount equal to the quality gap between the two network products: $p_R^2 < p_I^2 - [v_I(b_I + 2) - v_R(b_R + 1)]$. Bertrand competition results in the incumbent serving B_2 at a price $p_I^{*2} = c + v_I(b_I + 2) - v_R(b_R + 1)$.

If, instead, B_1 chose the rival, from assumption (1.21) and from $v'_I(n_i) \geq 0$, it follows that for the second buyer the quality of the rival's

network is superior to the incumbent's:

$$v_R(b_R + 2) > v_I(b_I + 2) \geq v_I(b_I + 1) \quad (1.24)$$

In this case, it is the incumbent that suffers a competitive disadvantage and must offer a discount in order to attract B_2 : $p_I^2 < p_R^2 - [v_R(b_R + 2) - v_I(b_I + 1)]$. In equilibrium, the rival supplies the second buyer at a price $p_R^{*2} = c + v_R(b_R + 2) - v_I(b_I + 1)$.

Let us move to the first period. Agents anticipate that the second buyer will follow the choice of the first one. Hence, B_1 is willing to address the incumbent if (and only if) $v_I(b_I + 2) - p_I^1 > v_R(b_R + 2) - p_R^1$. By assumption (1.21), at full size the rival's network exhibits higher quality than the incumbent's. This represents a disadvantage for the incumbent when competing for B_1 and calls for a discount relative to firm R 's price in order to win B_1 : $p_I^1 < p_R^1 - [v_R(b_R + 2) - v_I(b_I + 2)]$. However, the supplier who wins the first buyer will win also the second, thereby obtaining a total profit equal to:

$$\pi_i = p_i^1 + p_i^{*2} - 2c \quad (1.25)$$

with $i = I, R$. We can thus denote as $\tilde{c}_i = 2c - p_i^{*2} = c - [v_i(b_i + 2) - v_j(b_j + 1)]$ with $i \neq j = I, R$ each firm's adjusted cost to supply the first buyer, which corresponds to the total cost to supply the two buyers minus the rents extracted from the second one. Note that, even though higher quality at full size favours rent extraction by the rival, the fact that one buyer is insufficient for firm R to achieve efficient scale is favourable to the incumbent. If the latter effect is sufficiently strong, the incumbent extracts more rents than the rival from the second buyer and may manage to win the first buyer despite the discount it has to offer. This is the case if (and only if):

$$\tilde{c}_I < \tilde{c}_R - [v_R(b_R + 2) - v_I(b_I + 2)] \quad (1.26)$$

which is equivalent to

$$v_R(b_R + 2) < v_I(b_I + 2) + \frac{v_I(b_I + 1) - v_R(b_R + 1)}{2} \equiv v_P \quad (1.27)$$

with $v_P > v_I(b_I + 2)$ by assumption (1.22). From the equation above, it is clear that v_P is increasing in b_I .

It follows that when $v_R(b_R + 2) < v_P$, the incumbent wins B_1 and sells at a price $p_I^{*1} = \tilde{c}_R - [v_R(b_R + 2) - v_I(b_I + 2)] = c - [v_R(b_R + 2) - v_I(b_I + 1)] - [v_R(b_R + 2) - v_I(b_I + 2)] < c$ by assumption (1.21). If instead

$v_R(b_R + 2) \geq v_P$, then firm R secures B_1 and sells at a price $p_R^{*1} = \tilde{c}_I + [v_R(b_R + 2) - v_I(b_I + 2)] = c - [v_I(b_I + 2) - v_R(b_R + 1)] + [v_R(b_R + 2) - v_I(b_I + 2)]$. ■

Welfare analysis Under demand externality, the fact that, at full size, the quality of the rival's network is superior to the incumbent's is not sufficient to guarantee that exclusion of the rival is welfare-detrimental. Indeed, the utility of old consumers who bought from the incumbent in the past cannot be ignored when studying the welfare effects: even if they do not buy any longer, they continue using the network product and their utility is affected by the decision of the new customers.

To see whether predation is welfare-detrimental, let us compare total welfare at the exclusionary equilibrium with total welfare when the new buyers buy from the rival firm. Note that the assumption of inelastic demands implies that prices can be ignored when computing total welfare, as they reduce consumer surplus by the same amount as they increase profits.

Exclusion of the more efficient rival is welfare-detrimental if (and only if):

$$W^{exclusion} = b_R v_R(b_R) + (b_I + 2)v_I(b_I + 2) - 2c < (b_R + 2)v_R(b_R + 2) + b_I v_I(b_I) - 2c = W^{expansion}$$

Rearranging, exclusion is welfare-detrimental if the increase in the utility of the new customers if they joined the rival's network instead of the incumbent's (firm term on the left-hand side below) plus the increase in the utility of the rival's old customers if their network increased in size because of the new customers' decision to join it (second term on the left-hand side below) outweighs the increase in utility that the old customers of the incumbent enjoy when the new customers choose the incumbent's network:

$$2[v_R(b_R + 2) - v_I(b_I + 2)] + b_R[v_R(b_R + 2) - v_R(b_R)] > b_I[v_I(b_I + 2) - v_I(b_I)] \quad (1.28)$$

Note that when the utility functions are concave and reach an asymptote v at $n = b_I$, the fact that new customers join the incumbent's network does not increase the utility of the incumbent's old customers. Hence, the right-hand side in the above condition is zero, whereas the left-hand side is positive by assumption (1.21). Hence, in this specific case,

predation is welfare-detrimental when it arises at the equilibrium, that is, when $v_R(b_R + 2) < v_P = \frac{3v}{2} - \frac{v_R(b_R+1)}{2}$.

At the same time, one can find situations in which predation arises at the equilibrium (that is, $v_R(b_R + 2) < v_P$) and is not welfare-detrimental, precisely because of the role played by the externality on the old buyers.

1.3.2.8 Predation in two-sided markets*

In this section, we consider the case where each firm (or platform) can sell its product to two different groups of consumers, each group (or side of the market) benefiting from positive externalities from the number of users on the *other* side. We assume that a consumer on side k and using product i will receive a utility $U_{ki} = v_{ki}(n_{li}) - p_{ki}$, with $k, l = 1, 2, k \neq l, i = I, R$, with n_{li} being the total number of users (both old and new buyers) of platform i on side l and with $v'_{ki}(n_{li}) \geq 0$. Platforms are incompatible.

The incumbent and the rival have a constant unit cost c . Each platform has an installed base of old customers b_{ki} with $k = 1, 2, i = I, R$, who are not buying any longer, but continue to use the product. For simplicity, we assume that a given platform has the same customer base on each side: $b_{1I} = b_{2I} = b_I$ and $b_{1R} = b_{2R} = b_R$, with the *incumbency advantage* amounting to $b_I > b_R \geq 0$. We also assume that $v_{1i}(\cdot) = v_{2i}(\cdot) = v_i(\cdot)$, with $i = I, R$.

When the game starts, there are two *new* buyers, B_1 and B_2 , one on each side of the market, who are making purchase decisions sequentially.

Finally, similarly to the previous sections, we assume that the rival is overall more efficient but it has an initial disadvantage:

$$v_R(b_R + 1) > v_I(b_I + 1) \quad (1.29)$$

$$v_I(b_I) > v_R(b_R) \quad (1.30)$$

The game is the usual one, with firms first competing for B_1 and then for B_2 .

The following can be shown:

Proposition 1.10 *There exists a threshold level v'_P with $v'_P > v_I(b_I + 1)$ such that:*

- (**Predation**) *If $v_R(b_R + 1) < v'_P$, then the incumbent supplies both buyers. It sells below cost to the first buyer, while recouping on the second buyer: $p_I^{*1} = \tilde{c}_R - [v_R(b_R + 1) - v_I(b_I + 1)] < c$ and $p_I^{*2} = c + v_I(b_I + 1) - v_R(b_R) > c$.*

- **(Entry/Expansion)** If $v_R(b_R + 1) \geq v'_P$, then firm R supplies both buyers. The price paid by the first buyer is lower than the price paid by the second: $p_R^{*1} = \tilde{c}_I + [v_R(b_R + 1) - v_I(b_I + 1)] < c + v_R(b_R + 1) - v_I(b_I) = p_R^{*2}$.

The threshold v'_P is weakly increasing in b_I .

Proof. Proceed by backward induction and consider the second period. (a) If in the first period B_1 bought from I , then B_2 's utility from buying from I and from R , respectively, will be: $U_{2I} = v_I(b_I + 1) - p_I^2$ and $U_{2R} = v_R(b_R) - p_R^2$. Note that B_2 enjoys the additional benefit from one extra user on side-1 if she buys from I , but not from R . From assumption (1.30) and from $v_I(n_i)$ being (weakly) increasing in the total number of users, it follows that in order to attract B_2 the rival must offer a sufficiently large discount as compared to the incumbent's price: $p_R^2 < p_I^2 - [v_I(b_I + 1) - v_R(b_R)]$. Bertrand competition results in the incumbent serving B_2 at a price $p_I^{*2} = c + v_I(b_I + 1) - v_R(b_R)$. (b) If in the first period B_1 bought from R , then B_2 's utility from buying from I and from R , respectively, will be: $U_{2I} = v_I(b_I) - p_I^2$ and $U_{2R} = v_R(b_R + 1) - p_R^2$. This time, B_2 enjoys the additional benefit from one extra user on side-1 if she buys from R . From assumption (1.29) and from $v'_I(n_i) \geq 0$, it follows that it is the incumbent that suffers a competitive disadvantage and must offer a discount to attract B_2 : $p_I^2 < p_R^2 - [v_R(b_R + 1) - v_I(b_I)]$. In equilibrium, the rival supplies B_2 at a price $p_R^{*2} = c + v_R(b_R + 1) - v_I(b_I)$.

Consider now competition for B_1 . Agents anticipate that the second buyer will follow the choice of the first one. Hence, B_1 is willing to buy from the incumbent if (and only if) $v_I(b_I + 1) - p_I^1 > v_R(b_R + 1) - p_R^1$. By assumption (1.29), overall efficiency represents an advantage for firm R when competing for B_1 and the incumbent must offer a discount relative to firm R 's price in order to win B_1 : $p_I^1 < p_R^1 - [v_R(b_R + 1) - v_I(b_I + 1)]$. However, the platform that serves the side-1 buyer will also serve the side-2 buyer, thereby making total profits $\pi_i = p_i^1 + p_i^{*2} - 2c$, with $i = I, R$. Also in this case, we can denote as $\tilde{c}_i = 2c - p_i^{*2} = c - [v_i(b_i + 1) - v_j(b_j)]$, with $i \neq j = I, R$, each firm's adjusted cost to supply the first buyer. Again, higher overall efficiency favours the rival, but the initial advantage is favourable to the incumbent. If the latter effect is sufficiently strong, the incumbent extracts more rents than the rival from the second buyer and may manage to win the first buyer despite the discount it has

to offer. This is the case if (and only if):

$$\tilde{c}_I < \tilde{c}_R - [v_R(b_R + 1) - v_I(b_I + 1)] \quad (1.31)$$

which is equivalent to

$$v_R(b_R + 1) < v_I(b_I + 1) + \frac{v_I(b_I) - v_R(b_R)}{2} \equiv v'_P \quad (1.32)$$

with $v'_P > v_I(b_I + 1)$ by assumption (1.30). From the equation above, it is clear that v'_P is increasing in b_I .

Then, when $v_R(b_R + 1) < v'_P$, platform I wins competition for B_1 and sells at a price $p_I^{*1} = \tilde{c}_R - [v_R(b_R + 1) - v_I(b_I + 1)] = c - [v_R(b_R + 1) - v_I(b_I)] - [v_R(b_R + 1) - v_I(b_I + 1)] < c$ by assumption (1.29). When instead $v_R(b_R + 1) \geq v'_P$ it will be platform R which obtains B_1 , with $p_R^{*1} = \tilde{c}_I + [v_R(b_R + 1) - v_I(b_I + 1)] = c - [v_I(b_I + 1) - v_R(b_R)] + [v_R(b_R + 1) - v_I(b_I + 1)]$. ■

An application of this model can be used to rationalise the *Napp* case that we will discuss in Section 1.6.1. In that case, firms were selling to hospitals (our side-1) and to the community segment (side-2). While hospitals' utility was not influenced by decisions in the community segment, community decisions were heavily affected by hospitals'. In terms of our model, we would have $v_{1i}(\cdot) = \bar{v}_i$ while $v_{2i}(\cdot) = v_{2i}(n_{1i})$ – that is, utility of side-2 buyers (the community segment) depends on the number of side-1 buyers (the hospital segment).

Welfare analysis

The welfare analysis in the case of two-sided markets is quite similar to the one developed for the case of network externalities.

Since the two sides are perfectly symmetric, we develop the welfare analysis by looking at only one side of the market. Exclusion of the more efficient rival is welfare-detrimental if (and only if):

$$\begin{aligned} W^{exclusion} &= b_R v_R(b_R) + (b_I + 1)v_I(b_I + 1) - c < (b_R + 1)v_R(b_R + 1) \\ &+ b_I v_I(b_I) - c = W^{expansion} \end{aligned}$$

Rearranging, exclusion is welfare-detrimental if the increase in the utility of a new customer when she joins the rival's platform instead of the incumbent's (first term on the left-hand side below) and the increase

in the utility of the rival's old customers when their platform increases in size because of the new customers' decision to join it (second term on the left-hand side below) is large enough to dominate the increase in utility that the old customers of the incumbent would have enjoyed had the new customer chosen the incumbent's platform:

$$\begin{aligned} & [v_R(b_R + 1) - v_I(b_I + 1)] + b_R[v_R(b_R + 1) - v_R(b_R)] \\ & > b_I[v_I(b_I + 1) - v_I(b_I)] \end{aligned} \quad (1.33)$$

Note that when the utility functions are concave and reach an asymptote v at $n = b_I$, the fact that new customers join the incumbent's network does not increase the utility of the incumbent's old customers. Hence, the right-hand side in the above condition is zero, whereas the left-hand side is positive by assumption (1.29). Hence, predation is welfare-detrimental when it arises at the equilibrium, that is, when $v_R(b_R + 1) < v'_p = \frac{3v}{2} - \frac{v_R(b_R)}{2}$.

At the same time, one can find situations in which predation arises at the equilibrium (that is, $v_R(b_R + 1) < v'_p$) and is not welfare-detrimental, precisely because of the role played by the externality on the old buyers.

1.4 From Theory to Practice

In this section, we discuss the policy implications that can be drawn from the literature reviewed in this chapter and we identify some criteria that can guide antitrust intervention in the area of predatory allegations. To do so, we shall discuss the main implications of the theories of predation, which will lead us to propose a two-tier test requiring first the finding of dominance (as a proxy for recoupment) and then the adoption of a price-cost test (to show profit sacrifice). In what follows we will also highlight that the implementation of both steps should be far from mechanical and, instead, should be guided by a clear and consistent theory of harm.

1.4.1 Policy Implications from Theories of Predation: Profit Sacrifice and Recoupment

In the previous sections we have reviewed different theories of predatory pricing. The role of such theories should not only be to demonstrate that predation may occur in general, but also to help competition authorities

and courts identify the circumstances under which predation may take place in practice.

As we have seen, there are different mechanisms whereby an incumbent firm might profitably use a predatory strategy to deter entry of new competitors or to force exit of existing ones. However, whether relying on reputation, signalling, deep-pockets or denial of scale economies, theories of predatory pricing present some common pattern: at the beginning, during the predatory phase, there is a *sacrifice of profits*, which is then followed – after the rival has exited the market or has been marginalised, or the potential entrant has given up on entering – by a period during which there is *recoupment*, through high prices and profits.

Hence, we propose a two-tier test of predation.⁴⁹

First, *in lieu* of proving recoupment, it should be demonstrated that the alleged predator enjoys a strong dominant position.^{50,51} This is because only if a firm has a lot of market power, will it be able to recoup any profits lost during the predatory phase. At this stage spelling out a theory of harm, that is, looking for a coherent economic rationale behind the predatory conduct is fundamental because it helps understand the extent to which the alleged predator is able to recoup any profits lost during the predatory phase.

We do not believe, however, that the recoupment test should be as stringent as indicated by the US Supreme Court in its recent case-law (see Section 1.5.1), because this would make it almost impossible for plaintiffs to prove predation. And after all, a prerequisite for the finding of dominance is that the firm has substantial market power, which implies that current rivals do not exercise strong constraints on prices, and that sufficiently large barriers to entry exist in the industry, which makes it likely that – after the exit of the prey – the dominant firm will increase prices.

Second, we propose a particular way of interpreting the existence of profit sacrifice; that is, that there has been an *actual loss* by the alleged predator. This entails the use of a price-cost test, as will be discussed in more detail in Section 1.4.3. The spelling out of a coherent theory of harm

⁴⁹ Joskow and Klevoric (1979) were probably the first to suggest a two-tier test for predatory pricing. This type of tests has later been endorsed by many authors. Here we follow in particular Motta (2004: 442–3).

⁵⁰ The discussion of how to establish dominance in practice is beyond the scope of this book. On this, see for instance, Motta (2004: Chapter 3).

⁵¹ We comment below on why it would not be recommendable to pursue predatory cases against ‘borderline’ dominant firms.

provides crucial guidance also in the implementation of the price-cost test, as we will discuss below.

Note that in this two-tier test the analysis of dominance should come before the analysis of profit sacrifice. This is because the former will in most cases be simpler to carry out than the latter, and would allow to screen out some cases (those where the alleged predator is not dominant) saving the competition authorities and courts time and resources.

1.4.2 How to Recognise Predation: The Theory of Harm

As just mentioned above, the spelling out of a coherent theory of harm is fundamental both to understand the scope of recoupment and to implement the price-cost test. Of course, it is not sufficient to invoke a certain theory of harm, but it must be shown that the facts of the case are consistent with such a theory. We now turn to a discussion of observable factors that would be consistent with each theory. We first deal with the more traditional theories,⁵² and then with the scale economies theory that we presented in the previous sections.

1.4.2.1 Facts consistent with traditional theories of harm

In the case of a theory of harm based on *financial market predation* (see Section 1.2.2.1), there are two necessary features to make the predatory claim compelling in a real-world case.

The first step requires showing that *financial frictions* are particularly severe in the relevant industry. For instance, access to external finance tends to be problematic in industries with a low proportion of tangible assets, as assets that are more tangible increase the value that can be pledged to creditors in default states.⁵³ Likewise, financial constraints are more of a concern in innovative industries, as various factors limit the payouts that can be credibly pledged to external financiers, including the fact that informational asymmetries between managers and outside investors tend to be larger than in more mature industries.⁵⁴ In addition, financial frictions vary across countries, both due to historical and institutional

⁵² See also Bolton et al. (2000) for a detailed analysis and examples.

⁵³ The corporate finance literature has exploited in different ways the idea that tangible assets can reduce the severity of financial constraints. See, for instance, Almeida and Campello (2007) and Braun and Larrain (2005).

⁵⁴ Several papers have provided indirect evidence of severe financial constraints in innovative industries, by examining the sensitivity of R&D investment to cash flow shocks (see Hall, 2010, for a comprehensive survey). More recent evidence relies instead on firms' own assessment of financial constraints: Hajivassiliou and Savignac (2016) document that in

reasons, as well as to the development and the market structure of the financial industry. Institutional reasons include the quality of corporate governance regulations, which affect the scope for divergence between managers' behaviour and shareholders' interests. Access to finance also depends on the time period in question; for example, external financing becomes harder during or just after an economic or financial crisis, *ceteris paribus*, whereas booms are usually associated with looser credit conditions.

The second step requires showing that there is an *asymmetry between the predator and the prey* in terms of ability to raise external funds, for instance because the predator owns more collateralisable assets, can rely on more (internally generated) liquidity or is part of a conglomerate with an active internal capital market.⁵⁵ Note that this asymmetry may make new or recent entrants vulnerable to predatory strategies particularly in industries that for intrinsic (technological) reasons depend on external finance more than others. For instance, firms (especially those in their early life stage) are highly dependent on external sources of finance in industries that require a large initial scale, continuing investment and have a large gestation period before cash flows are harvested (think of the pharmaceutical as opposed to the textile industry).⁵⁶

In this environment, it is more likely that predatory pricing would deprive the prey of the cash flow that it crucially needs to obtain external funding and continue operating.

In the case of a theory of harm based on *reputation effects*, the discussion in Section 1.2.2.2 has shown that the necessary ingredients are: (i) the existence of multiple markets (either different goods or different geographic markets) with potential entry in each case; (ii) the notion that aggressive behaviour at earlier stages (or in some markets) increases the probability perceived by rivals that the incumbent will be aggressive later (or in other markets).

French manufacturing industries innovative firms are more likely to report difficulties in raising external capital.

⁵⁵ A copious literature in corporate finance, dating back to Fazzari, Hubbard and Petersen (1988), has emphasized how the availability of internally generated cash affects firms' real investment decisions by alleviating their financial constraints. Recent empirical findings in Maksimovic and Phillips (2008), Kuppaswamy and Villalonga (2016), Almeida et al. (2016) and Boutin et al. (2013) show that multidivisional firms and business groups mitigate financial constraints faced by segments operating in industries where access to external funding is particularly problematic.

⁵⁶ A vast literature starting with Rajan and Zingales (1998) has highlighted that financial frictions have more significant real effects in financially dependent industries.

Assessing the expectations of the market agents and their interpretation of competitors' strategies is quite demanding. Some evidence (for example, internal documents of a corporate plan, press releases) that shows that the alleged predator consciously behaves aggressively in order to build such a reputation for entry-deterrence may be useful.

As for *signalling theories of predation* (see Section 1.2.2.3), we believe that they are of rather limited practical relevance. In the case of cost-signalling, for instance, one should look for observable actions (including prices) by the incumbent that make the potential entrant conjecture that the variable costs of the alleged predator are lower than they actually are. Furthermore, one should be able to prove that the incumbent would not have taken the same actions had it not faced entry. It is difficult to imagine that a competition authority may build a solid predation case on this basis.⁵⁷

By contrast, one of the advantages of the theory of predation we have proposed in Section 1.3 is that it allows us to specify a number of arguably more readily observable factors that make (profitable) predation more or less likely. We now turn to a discussion of such factors, which a competition authority referring to such theory in a given case should show to hold in the industry at issue.

1.4.2.2 Facts consistent with a scale-economies theory of harm

- *Economies of scale* – whether due to fixed costs, learning effects, demand externalities, two-sided market effects, or other reasons – are obviously crucial for this theory, where predation hinges upon depriving the rival of the scale it needs to achieve minimum efficient scale. Recall though that predation may be used not only to exclude a rival from the industry altogether, but also to relegate it to a market niche. So the fact that the rival has reached minimum efficient scale in a segment, region or niche of the market does not imply that scale economies do not play a role any longer: scale may be crucial

⁵⁷ Bolton et al. (2000) argue that signalling theories of predation may represent a plausible predatory theory of harm, and identify the *General Foods* case as a possible illustration of test-market predation. In this case, General Foods, owner of the Maxwell House coffee brand, sharply decreased its price when Folger coffee entered in the eastern states of the US. Hilke and Nelson (1987) rather interpreted the facts of this case in the spirit of a reputation-based theory of harm. The two stories are not observationally inconsistent, although the interpretation provided by Hilke and Nelson (who worked on the case as FTC economists) may reflect the facts of the case more closely in our view than one based on test-market predation. The FTC dismissed the case based on a finding that General Foods lacked monopoly power, in a market that the FTC defined more broadly than the plaintiff. The FTC also found barriers to entry and to expansion to be low.

to recover investments by the rival in additional capacity, to build a more connected network, to expand in new markets, to provide new products, and so on. Moreover, the scale-economies theory also fits situations in which success in one market is key to enter in another market.

- There must be a *strong incumbency advantage*, which, as a first approximation, may be captured by a very high and persistent market share of the incumbent; this would be reinforced by the presence of switching costs, by the infrequency of purchases and by demand externalities. In other words, the existence of a dominant position is an important requirement for a theory of harm based on the scale-economies mechanism, which reinforces the role of dominance as a necessary prerequisite for a finding of predation.

However, it would make little sense to open a predatory case when there is borderline dominance. Indeed, recall that the higher the proportion of captive buyers (or the larger the established base) of the incumbent relative to the rival, the more likely for predation to occur. Therefore, the theory suggests that a predation case where the defendant has 70–80 per cent of the market will be (other things being equal) much stronger than one where the defendant has 40 per cent of the market. In some jurisdictions, it is possible to find that a firm has a dominant position even with a relatively low market share (for instance, in South Africa the Competition Act defines as dominant any firm with a market share above 35 per cent), but the predation theory based on scale economies and incumbency advantage suggests that it makes sense to potentially be concerned only where the alleged predator has a really strong dominant position.

- By its very nature, predation consists of low prices, followed by high prices. Quite trivially, therefore, the possibility to engage into *intertemporal price discrimination* is another necessary condition for predation. If for any reason there are obstacles to price discrimination over time, then predation will not be possible.
- *Weak buyer power* (and uncoordinated buyers) is another crucial determinant of the theory. Indeed, predation hinges upon the ability to exploit the externality that earlier buyers will impose upon later ones, and this is possible only to the extent that buyers cannot take joint decisions or cannot pool their orders into a single period. Furthermore, the existence of large buyers, whose demand is sufficient for the entrant to reach efficient scale, limits or indeed removes the scope for predation.

- Under this theory, predation is more likely to be a potential concern in *mature markets*: first, a rapidly growing market is one where demand in later periods is likely to be sufficient for the prey to reach minimum efficient scale, thereby making predation on early buyers unfeasible; second, the number of contestable buyers will be larger relative to the captive ones, limiting the incumbency advantage. As a result, we believe that predation based on the scale-economies mechanism presented in this chapter is less likely to occur in industries where there are no entrenched market positions or where there is a rapidly expanding demand.
- Note, however, there are other reasons why predation may be a concern in *markets that are expected to grow considerably* in the future or that may undergo an evolutionary change. Consider for instance a market in which a new technology, which represents an alternative to the existing technology dominated by the incumbent, can be introduced. In such a context, the entrant may suffer from both the need to win sufficient scale to challenge the incumbent on the whole market, and the need to receive sufficient backing from outside investors. Note that financial frictions may be particularly severe in a period when the prospects of the new technology are difficult to assess. This creates the scope for financial predation by the firm that dominates the old technology: by setting aggressive prices in early periods, it may prevent the rival from obtaining the funds that are key to invest and compete in later periods when demand for the new technology will grow.⁵⁸ Similarly, the dominant firm may have an incentive to be aggressive on early buyers because they are the crucial ones to legitimise the entrant in the eyes of external financiers and to facilitate it in obtaining funds for future investments. In this case, even though the demand of early buyers is small compared to future demand, it may still be critical for the expansion of the new rival.
- We have also seen that predation may be more of a potential concern, all else equal, if there is *weak downstream competition*: if buyers are downstream firms (for example, retailers) which compete fiercely for final consumers, few buyers will be able to bring large business scale to a supplier, and hence it will be more difficult for the dominant

⁵⁸ Of course, this argument hinges upon the assumption that dominance over the old technology makes the incumbent less vulnerable to financial constraints than the entrant. Should the entrant be an established firm coming from a different sector, it may well enjoy enough assets for obtaining credit not to be an issue.

incumbent to prevent the rival from reaching the scale it needs (in the limit, the incumbent would have to ‘block’ all the buyers because any of them could bring sufficient orders to the prey: but then, no recoupment would be possible). Clearly, the degree of competition is not a variable which can be easily measured. However, a careful analysis of the industry may give hints in this respect. For instance, an industry where firms sell products that are highly differentiated (either physically or geographically), or where capacity constraints are prevalent is an industry where product market competition is likely to be weaker than otherwise. Similarly, we expect downstream competition to be weaker – other things being equal – in industries where switching costs and lock-in effects are important.

1.4.3 How to Recognise Predation: Tests for Profit Sacrifice

The second tier of the test involves the establishment of a sacrifice of profits. Here, the crucial question is what exactly is meant by *sacrifice of profits*. Indeed, in the traditional (financial, reputation, signalling) theories of predation, sacrifice of profits does not necessarily mean that the predator incurs actual losses; it may simply mean that it earns lower profits than it would have earned if it had behaved in a non-abusive way. This opens the possibility to find predation also in cases where prices are above some relevant measures of cost.⁵⁹

In principle, therefore, a literal interpretation of economic theory would require a comparison between actual (allegedly predatory) prices and some counterfactual optimal prices that the incumbent would set if it accommodated the rival. However, the calculation of such counterfactual prices would be extremely difficult, and would probably give rise to mistakes and huge uncertainty in the process.⁶⁰

⁵⁹ Under the scale economies theory, there is in principle similar uncertainty over how to define ‘profit sacrifice’ in practice. Indeed, we have seen in Section 1.3.2.2 that when exclusion occurs because of buyers’ coordination failures, i.e. when offers are made simultaneously to all buyers, below-cost pricing may not occur.

⁶⁰ Ordovery and Willig (1981) propose a test whereby an action is deemed predatory if it is optimal only because it induces the exit of a rival. Similarly, according to the ‘no economic sense’ test a conduct is considered as *exclusionary* if it ‘would make no economic sense for the defendant but for the tendency to eliminate or lessen competition’. Indeed, some versions of these tests have often been proposed to identify not only predatory pricing but exclusionary practices in general. We find it difficult to use a general ‘test’ of this type to assess conducts which are very different (for instance, exclusive dealing may be exclusionary but it does not necessarily entail profit sacrifice), and we believe that they are

Therefore, we suggest to interpret ‘profit sacrifice’ in a stricter way, that is, as ‘negative profits’, and to use an Areeda-Turner’s type price-cost test (see Areeda and Turner, 1975). The logic of this test is to identify two cost benchmarks, an *upper bound* and a *lower bound*. (We shall discuss the cost thresholds below, but for the time being one can associate the upper bound with Average Total Costs, ‘ATC’ and the lower bound with Average Variable Costs, ‘AVC’).

If the dominant firm’s prices are *above* the upper bound, then predatory allegations should be dismissed. Hence, prices above the upper bound represent a ‘safe harbour’ for dominant firms.

Instead, if prices are *below* the lower bound, there should be a strong presumption of abusive conduct, with the defendant having the burden of proving otherwise, in particular that there are objective justifications for below-cost pricing.

Finally, with prices falling *between* these two thresholds, the plaintiff (or the competition authority) should have the burden of proving that predation took place. Similarly, there may be cases in which the evidence is uncertain, for instance because there may be different methods of estimating cost thresholds. We would treat such cases like those in which prices fall between the two thresholds. Under these circumstances, spelling out a rigorous theory of harm – that is, showing that there is a precise theory that rationalises the predatory conduct and that the facts of the case are consistent with that theory – plays an even more pivotal role than otherwise.

Documentary evidence proving the *intent* to exclude may also play an important role when prices fall between the two thresholds. From an economic point of view, evidence from internal documents that the alleged predator’s managers would be happy if competitors exited the market is not by itself relevant: after all, this is what one should expect to find in any situation where firms are fiercely competing. However, if there was evidence showing that a coherent business strategy has been put in place with exclusionary purposes, and especially if those documents reveal the intention to sacrifice profits to achieve that aim, and this was consistent with the theory of harm put forward by the plaintiff, this would certainly represent strong evidence of predatory conduct. This would be the case if, for instance, in an environment in which financial frictions are severe, a competition authority found the existence of internal documents of the

very difficult to administer in practice. We refer the interested reader to the discussions in Vickers (2005), Kobayashi (2010), Salop (2006) and Werden (2006).

alleged predator showing particular attention to the financial situation of the prey and indicating a rational strategy to limit significantly the prey's access to external funding. This approach is in line with the EU law, according to which the existence of evidence pointing to a strategy to exclude a competitor may help determine the illegality of a conduct in a case where prices are below average total costs but above average variable costs (see Section 1.5.2).

The process of assessing whether the facts of the case are consistent with the proposed theory of harm may also involve the analysis of the *effects* of the alleged predatory conduct on the competitors of the dominant firm. The observation that their market shares are stable or increase during the allegedly predatory phase, for instance, may be hard to reconcile with a coherent predatory conduct undertaken by the dominant firm.⁶¹

Consider for instance a market in which there is a dominant firm. Suppose there is entry by a very large multinational firm endowed with a lot of financial resources, whose products are internationally known and well regarded, and which has entered the market with a very large, modern and efficient production plant, and a very large marketing campaign, resulting in a significant market share gained in little time, unaffected (both financially and in its market performance) by aggressive price counter-moves by the dominant firm. We might even find that some of the dominant firm's brands are sold below one or the other cost threshold, but the lack of effects on the rival would beg some questions. Perhaps the strong price reaction is an attempt by the dominant firm not to lose too much market share too quickly, and since it is unlikely that the large rival would exit (given the financial resources available and the strong commitment in production and brand recognition), the facts of the case would hardly point to a predatory strategy. Further, an intervention by the antitrust authority may completely backfire, converting what looks like healthy competition between two powerful firms wrestling for domination of the market into possible soft competition, induced by the fear of the dominant firm of possible antitrust penalties, and accommodated by the powerful entrant which once obtained a large share of the market may prefer a more accommodating pricing strategy.

⁶¹ This does not mean, of course, that in order to find predation we should always observe market exit or a strong downsizing in the market share of the rival of the dominant firm. On some occasions, the prey may be willing to undertake losses not to lose customer base and perhaps while waiting for antitrust intervention.

1.4.3.1 Too favourable to defendants?

The price-cost test described above is admittedly favourable to the defendant. First, because it may lead to some false negatives (type-II errors): profit sacrifice does not necessarily mean actual losses, so there may be cases where predation *did* take place but would not be detected by the test. Second, because as we will discuss in the next sections, showing that prices are below some relevant measures of costs is a tricky exercise. As a consequence, because of lack of data, or because of different classifications of costs, a cautious agency may be unable to determine with reasonable certainty that prices are below costs.

However, the price-cost test is easier to administer than a rule that requires actual profits to be compared with the hypothetical profits that the dominant firm would have made absent the exclusionary motive.⁶²

The price-cost test also limits the risk of stifling price competition by decreasing the rate of false positives (type-I errors). Low prices are precisely what one would like to have from a consumer welfare perspective, and the outcome of competitive markets. Distinguishing bad (predatory) low prices from good (competitive) low prices is extremely difficult, and it is crucial to avoid rules which – to prevent the former – end up discouraging the latter.⁶³

For these reasons, and most notably for the sake of administrability, we suggest treating prices above the upper bound of costs as a safe harbour.

However, we would not consider it outrageous if a judge or a competition authority decided to depart from this approach in *exceptional circumstances*. If, for instance, the antitrust agency found that the dominant firm is pricing above (the upper bound of) costs, but it held strong documentary evidence of predatory conduct (for instance, showing that the top management had deliberately followed a certain strategy which leads to

⁶² For the importance of administrability when setting policy on predatory pricing (and more generally in antitrust), see (then judge) Stephen Breyer in a classic quote from *Barry Wright*: ‘[...]while technical economic discussion helps to inform the antitrust laws, those laws cannot precisely replicate the economists’ (sometimes conflicting) views. For, unlike economics, law is an administrative system the effects of which depend upon the content of rules and precedents only as they are applied by judges and juries in courts and by lawyers advising their clients. Rules that seek to embody every economic complexity and qualification may well, through the vagaries of administration, prove counter-productive, undercutting the very economic ends they seek to serve.’ Also for this reason, Breyer argues that although it may conceivably be anti-competitive in very particular situations, it is best to consider prices above average total costs as a safe harbour.

⁶³ See also Elhauge (2003a) for a discussion of restrictions to above-cost pricing. He concludes against such restrictions.

lower profits for the sole purpose of excluding a rival), then we would find it difficult to dismiss the allegation of abusive behaviour.

Furthermore, using above cost-pricing as a safe harbour reflects great concerns about over-enforcement and the risk of chilling competition. This makes a lot of sense in those countries where there is sufficient ease of entry (financial markets work reasonably well, administrative barriers to operate are low) and one can therefore rely on entry as a disciplining device. However, in less developed countries, where barriers to entry tend to be higher, markets are sometimes smaller, and entrenched dominant positions often prevail in many sectors, the major concern should probably be about under-enforcement. In that case, it would appear legitimate to us for competition authorities and courts in such countries to have a less conservative approach, and for instance adopt a presumption of abusive behaviour whenever prices are below ATC, or intervene also when a small proportion of the market is affected by the alleged exclusionary conduct (see discussion below).

1.4.3.2 When are prices below costs legitimate?

The test described above raises the question of whether circumstances may exist in which an alleged predator is legitimately setting below-cost prices.

In general, there may be many reasons why firms might price below costs as part of a normal competitive process (think of product complementarities, of the need to enhance consumers' goodwill, of markets characterised by switching costs, learning economies, network effects, or two-sided externalities: in all these cases, one would expect firms to sell below cost, absent any intention to exclude).

In most cases, however, such arguments would *not* apply to a firm which is already dominant in the relevant market where the alleged predation takes place. Dominance implies that the firm is already well established in the market: therefore, a dominant firm does not need promotional pricing to make itself known, nor does it need to lower prices to win customers locked in by switching costs or network effects, nor does it need to expand production to benefit from learning effects since it has probably already reached the minimum efficient scale of production and benefited from learning effects. For the same reason, a meeting-competition defence argument to justify below-cost pricing by a dominant firm should not be accepted.⁶⁴

⁶⁴ This principle is well established in EU law. See, for instance, *France Télécom*.

In other cases, though, one cannot immediately dismiss the possibility that below-cost pricing is pro-competitive or in any case unrelated to a predatory strategy. Consider for instance a firm selling complementary products, or a firm selling to two sides of the market. In these cases, pricing one product below-cost on one side of the market may be a standard business practice. It may be possible, for instance, that firms in the industry have always used this pricing structure, and this should obviously be taken into account: one cannot ask a dominant firm to raise the price of razors (or to charge a positive price for its newspaper) if the business model has always been to sell razors below cost and blades at a higher price (or to give away newspapers for free while charging advertisers). In those cases, though, it is important to assess whether the below-cost pricing has always been adopted or it has been introduced only when rivals start to threaten the firm's dominant position. For instance, in Section 1.6.1 we shall see how the UK Competition Appeal Tribunal did not accept Napp's defence that below-cost pricing was motivated by its selling in a two-sided type market when it found that losses were made only for those product specifications in which Napp was facing competitors.

Likewise, below-cost pricing may be accepted as a defence if an industry features excess capacity and is currently downsizing, as operating plants at higher utilisation levels can be efficient.

Another circumstance in which a dominant firm may be justified in its selling below-cost may be when a new product or standard appears in the market. Consider for instance internet access services. At a certain point in time, telecom operators upgraded their fixed networks and began offering retail broadband access services, which offered better (faster) internet connections to residential and business users; these thus started to replace (slower) narrowband access services. Shortly after broadband services became available, it would have probably made sense to define the relevant market as the market for all (narrowband and broadband) retail internet access services. This would have implied that an operator that was historically dominant in retail narrowband access services would have also likely been dominant in this new combined market at its infancy. However, one has to be cautious if wishing to prevent the dominant operator in the old technology from offering good deals (which may include promotional below-cost prices) in the new technology. Such promotional pricing may well have the effect of stimulating demand, operating closer to minimum efficient scale, and promoting the use of complementary services (for example, live TV streaming, online gaming, etc.) which use the band intensively. Note also that this is consistent with our warning against

predatory cases when there is a growing market and highlights again the importance of checking whether the facts of the case are consistent with the proposed theory of harm.

Furthermore, a predatory strategy should be systematic and not occasional: the finding that a dominant firm has charged below-cost prices only in very precise and limited circumstances (for instance, to dispose of existing stock or perishable goods) or to a very small part of the market, or on a small volume of orders should generally not be sufficient to support a predatory case. For instance, in terms of our theory of predation, it would be difficult to argue that a rival is denied sufficient scale when only – say – 5 per cent of the market has been affected by the predatory episode. However, as we stressed in the previous section, one should take into account that in some cases demand in early periods, even though small as compared to future demand, may still be crucial for a new firm to be able to invest and develop in the future. Hence, predation on (relatively) small volumes may still be a successful strategy to exclude the rival. This observation highlights again the importance of spelling out a rigorous theory of harm to interpret the facts of the case and to avoid a mechanical implementation of the test.

1.4.3.3 Cost benchmarks

The cost benchmarks that have been most often used in practice when assessing allegations of predatory pricing are the *average total costs* ('ATC') of the dominant firm – as an upper bound – and its *average variable costs* ('AVC') – as a lower bound. Variable costs are those that increase with the amount of output produced (for example, energy or labour costs, in most settings). Fixed costs, by contrast, do not vary as output increases (for example, management's office rent). The term 'average' (or 'unit') refers to the fact that costs incurred over the whole output level are divided by the level of output produced.

While these definitions seem fairly straightforward, following this taxonomy in practice may be more difficult. For instance, advertising costs are typically categorised as fixed. But we will see in Section 1.6 how, in *Wanadoo*, the European Commission held that France Télécom's marketing costs incurred in promoting its new broadband products in 2001 should be deemed to be variable, on the basis that France Télécom incurred 'advertising and marketing [costs] specific to Wanadoo Interactive's ADSL services' and that these were 'quite specific and correspond[ed] to successive campaigns which focus[ed] strongly on sales promotion.'⁶⁵

⁶⁵ See paras 62 and 64 of the Commission's decision.

In *Aberdeen Journals* (which we discuss in Section 1.6) the Office of Fair Trading ('OFT', the then UK competition authority) found that predation in the market for advertising space in local newspapers occurred only over one month, March 2000.⁶⁶ Over such a short timeframe, the OFT recognised that most costs would be fixed and very few costs would be variable. In doing so, the OFT only considered the newsprint and circulation costs as variable. The costs of the editorial staff and the advertising team were instead considered to be fixed given the very short timeframe considered.⁶⁷ (See below for further discussion on the impact of the time dimension on predatory cases.)

Partly to overcome the difficulty in determining the fixed or variable nature of a certain cost, the notion of *long run average incremental costs* ('LRAIC') has been proposed as a substitute for ATC.⁶⁸ This concept is particularly well suited to the context of firms producing multiple products, which usually face common costs, that is, (typically fixed) costs incurred in the production of more than one good or in the provision of more than one service (for example, a fixed telecoms operator historically used a copper network both for voice services and for broadband). The LRAIC of a product is the firm's total costs (when also producing the good at issue) minus what the firm's total cost base would have been had it not produced that good, divided by the quantity of the good produced. LRAIC thus includes all product-specific costs incurred in the research, development and marketing of the allegedly predatory product, even if those costs were incurred before the alleged period of predatory pricing. As Bolton et al. (2000) stress, LRAIC is a superior cost measure over ATC for a firm producing multiple products because it does not require joint and common costs (these are typically excluded from LRAIC) to be allocated, an exercise which lacks a precise methodology and is particularly unsuited for jury resolution (or for decision-making by a competition authority lacking the resources and expertise of a sector regulator).

For instance, the European Commission opted for an 'average incremental cost' benchmark in *Deutsche Post* (which we discuss in more detail

⁶⁶ A previous decision by the OFT in 2001 had been set aside by Competition Appeal Tribunal due to an unsatisfactory definition of the relevant market. The OFT could not condemn the abusive practice over an earlier period as the Competition Act 1998 only entered into force on 1 March 2000.

⁶⁷ The Competition Appeal Tribunal would have broadened the scope of the variable costs (and included further printing costs), but it did not overrule the OFT on the basis that it treated the evidence in the most favourable way to the defendant.

⁶⁸ The original proposal is due to Joskow and Klevorick (1979). See also Bolton et al. (2000).

in Section 1.6). This case focused on the German market for mail order parcel services. The Commission explicitly excluded all common costs that would be incurred upon the provision of other postal services (such as the costs to build and maintain Deutsche Post's network of freight centres and delivery points, used for a variety of services). The Commission supported this with reference to the fact that Deutsche Post incurred costs to fulfil its statutory obligation to maintain network reserve capacity, but that these should not 'burden' its mail order parcel services, which was a liberalised market. By contrast, the Commission included in the relevant cost measure 'the staffing and equipment costs of [sorting] activities dependent on the volume processed' (para. 13 of the decision).

Note that we interpret the price-cost test as a test for profit sacrifice. Another frequent interpretation is that of a *replicability* test.⁶⁹ Following this logic, a case in which the dominant firm's prices are above ATC (or LRAIC) should be dismissed because a rival firm at least as efficient as the dominant firm can profitably match such prices. Thus, such a cost standard identifies a reasonable safe harbour for the pricing policy of the dominant firm.

Concerning the lower bound, there are two variants of AVC that may be considered. One is the notion of *short run average incremental costs* ('SRAIC'), which is the average of all the (short-run) costs associated with an increase in production, in this case the extra output due to the allegedly predatory conduct. The notion of SRAIC includes both fixed and variable costs, but does not include the costs incurred *before* the period in which the allegedly abusive conduct took place (for example, the costs for infrastructure set up before the predatory period and that is used for the incremental production associated with the predatory output). Such costs are instead included in the LRAIC.

The other is the notion of *average avoidable costs* ('AAC'), which is the average of all variable costs and fixed costs that are *not sunk* and that are associated with the predatory output, so that they can be avoided if the firm stops producing.⁷⁰ One of the key advantages of the AAC standard is that it does not call for a categorisation between fixed and variable costs, which can often be problematic as we saw above. One drawback of the AAC measure, on the other hand, is that it is not always straightforward

⁶⁹ We note that the tests we discuss in Chapter 2 (on rebates) and in Chapter 5 (on margin squeeze) can also be interpreted as profit sacrifice tests for the incumbent, or as tests of replicability by a rival which is at least as efficient as the incumbent.

⁷⁰ Quoting Baumol (1996), 'avoidable costs can be thought of as the *decremental* cost to a firm if it decides to stop production.'

to determine which assets are not sunk and therefore could be redeployed upon exit/production stop and at what value, especially in the context of an event that never took place, namely the incumbent's exit (see O'Donoghue and Padilla, 2013: 304ff). In *Cardiff Bus* (which we review in greater detail in Section 1.6), the OFT found AAC to be the most appropriate benchmark.⁷¹ Specifically, the OFT included drivers' payroll, fuel, tyres, lubricants, insurance, road tax and maintenance, as well as the costs of restoring certain buses to deploy them on the 'white service routes' (that is, where the OFT found predation). The OFT found that the revenues from these routes did not even cover the salaries of the drivers who were serving these routes.

Note that, in practice, AAC and SRAIC are often used as synonymous. Indeed, in many situations AAC and SRAIC coincide, but this is not always the case and failing to distinguish between them may create confusion. What matters in computing the AAC is that a given cost is not sunk; whereas what matters in computing SRAIC is that a given cost has not been incurred before the predatory period.

Moreover, from a replicability perspective, the choice between the two notions of costs depends on the specific case at hand.⁷² If the predatory strategy aims at taking demand away from an *existing* rival and inducing it to leave the market, the pertinent notion is the one of the AAC, because, if the price of the dominant firm is below the AAC, an as-efficient rival is better off leaving the market rather than making a counteroffer. If instead, the predatory strategy aims at excluding a competitor from a *new* market (think, for instance, of a procurement auction to win a given amount of production), the pertinent notion is the one of SRAIC. Indeed, a price below the SRAIC indicates that a firm is not covering both the variable costs and the additional investments made to increase production. Then an as-efficient competitor cannot serve the auctioned production without incurring a loss.⁷³

⁷¹ See paras 7.150 *et seq.* of the OFT's decision.

⁷² We are grateful to Michele Polo for sharing insights with us on the discussion that follows.

⁷³ Consider the following stylised example to see how SRAIC, may sometimes differ from AAC. Imagine that a large, well-established chain of hairdressers that has hitherto only catered for men begins to cater for women as well (who, for simplicity, are offered just one type of service). To do so, this firm needs to rent more salons, hire specialist hairdressers and pay an annual insurance premium (which does not depend on the number of salons or customers). It also needs to run a large marketing campaign, to let potential customers know that it is beginning to cater for women as well. After one year of operations, one may consider the following categories of costs for the women's hairdressing business.

The time dimension is crucial to determine which costs are fixed, variable, avoidable or incremental. Over a sufficiently long period of time, all costs will be variable. Indeed, the 'long run' in microeconomics is often defined as the period over which firms can adjust all factors of production (that is, all costs become variable). The UK Competition Appeal Tribunal recognised this:

The longer the period that is taken, the more likely it is that cost will be classified as variable since, for example, over a longer timescale, employees can be dismissed or plant closed in response to changes in output. (Para. 353)

The relevant timeframe has to be assessed on a case-by-case basis. In *Aberdeen Journals* the incumbent's behaviour (in the market for advertising space in local newspapers) was deemed predatory only with reference to the month of March 2000. The OFT had found – also by analysing the monthly management accounts available – that the revenues received from advertisers for Aberdeen Journals's *Herald & Post* paper did not cover its average variable costs of production over that month (note that this weekly attracted no revenues from readers, as it was distributed free of charge). In fact, in March 2000, the losses from the *Herald & Post* equalled 42 per cent of its revenues.

In *Deutsche Post*, the Commission considered the six-year period from 1990 to 1995. A paper reviewing this case (OFT, 2005) suggested that the Commission's rationale for adopting such a temporal horizon was as follows:

Over six years all overheads (ie: costs that on a short term basis do not vary directly with output) were deemed avoidable because this was ample time to realize that the mail order parcel service was not covering such costs and therefore raise prices or dispose of the additional assets required to provide that service. To put it another way, over this period [Deutsche Post] would have had the choice not to re-incur all incremental costs of running its mail order parcel services business. (Para. A.21)

The short run incremental costs are the costs of incremental 'production'. These therefore include the variable costs (salons rental and specialist hairdressers' salaries) and the fixed costs (the insurance premium plus the cost of the marketing campaign).

The avoidable costs, instead, are the costs incurred during the period that the firm could avoid by discontinuing the women's hairdressing business. Such costs therefore include the variable costs as defined above and the annual insurance premium. They do not include, however, the costs of the marketing campaign, since these are completely sunk.

For both categories, these costs can be divided by the number of customers (women) during the period to determine the equivalent 'average' measures.

In *Wanadoo*, the Commission held that the relevant timeframe over which the customer acquisition costs should be assessed was the average length of a subscription (48 months).

In *Cardiff Bus*, the OFT opted for a 10-month timeframe (19 April 2004 to 18 February 2005), corresponding to the period over which Cardiff Bus operated one or more of its ‘white service routes’ (that is, over which the OFT found predation).

1.5 Case-law

In this section we provide a brief account of the history and evolution of landmark decisions by competition authorities and of the key case-law on predation on both sides of the Atlantic, paying attention to key judgments and to how enforcement policy evolved, all the way to recent cases. This overview will highlight a clear divergence in the US and EU jurisprudence: while in the US in the last 20 years we hardly observe cases involving predatory pricing, in particular cases favourable to the plaintiff, in Europe we observe a number of recent cases where the prices set by a dominant firm have been considered abusive. We proceed chronologically, beginning with US cases, followed by EU ones.

1.5.1 United States

In the US, the complaints by small firms of unfair business practices (including predatory pricing) adopted by their large rivals, which allegedly aimed to drive them out of business, are one of the reasons that led to the adoption of the Sherman Act in 1890. These concerns were addressed in Section 2 of the Sherman Act, which prohibits monopolisation, attempts to monopolise and conspiracies to monopolise any part of the trade or commerce among the several states, or with foreign nations.

The monopolisation provisions in the Sherman Act (1890) were tested before the Supreme Court already in *Standard Oil* (1910) and *American Tobacco* (1911), two cases affecting different industries but characterised by very similar behaviour by the incumbents. Neither case was specifically about predatory pricing, but in both instances the Supreme Court found strong evidence of a broad set of exclusionary practices, aimed at threatening and then eliminating competitors. The key behaviour in both cases was characterised by a sequence of forced acquisitions of competitors,

together with ancillary exclusionary conduct.⁷⁴ We describe these two cases in greater detail in Section 1.6.11, where we explain how the facts could be read in the context of a reputation effects' model.⁷⁵

More than half a century went before another notable case with a predatory element reached the Supreme Court. *Utah Pie* concerned the frozen (dessert) pies market, in the period from 1958 to 1961.⁷⁶ The *plaintiff* (not the defendant) had two-thirds of the market at the beginning of the predatory phase (this dropped to around 45 per cent at the end and allowed the plaintiff to earn positive profits). Yet in 1967 the Supreme Court ruled in favour of the plaintiff. This judgment was based, according to leading commentators, on evidence of 'occasional prices below full cost, weak evidence of animus against the plaintiff, and a declining price structure' (Areeda and Hovenkamp, 2006). Areeda and Hovenkamp also point to the presence of a competitive fringe of six or seven smaller rivals to observe that entry barriers could not be particularly high either, thus making it very difficult to believe that a predator could foresee sufficient recoupment.

Both economists and legal scholars remained puzzled by *Utah Pie*: in a market with relatively low barriers to entry and the presence of a competitive fringe, how could consumers be harmed if a firm with two thirds of the market loses 20 percentage points of market share following price competition by a rival? It is in this context that the scholarly article by Areeda and Turner was published in 1975. It immediately caused a structural break in how antitrust law was applied to predatory pricing cases, in the US and beyond. The basic idea of their test (similarly to what we discussed in Section 1.4) was to draw the line between abusive conduct and competition on the merits at the average variable cost level (a reasonable proxy for marginal cost): only prices below this line truly entailed a short-term profit sacrifice (an actual economic loss) and should thus be presumed unlawful (even if these prices are set to match those of a rival). Prices above average variable costs should be presumed lawful. Temporary promotional pricing, the authors added, should not be deemed harmful to the competitive process either.

⁷⁴ For a richer summary of the facts in these two cases, see Evans (1912).

⁷⁵ In addition to the reputation effect, these two cases could be read in the spirit of a signalling model based on 'predation-through-merger'.

⁷⁶ Here we also draw from Areeda and Hovenkamp (2006). This case was not assessed under the Sherman Act but under the Robinson-Patman Act (price discrimination leading to competitive injury).

Another important milestone in US antitrust law on predation is *William Inglis*, which dealt with the pricing of ‘white pan bread loaves’ in California between the mid-1960s and the mid-1970s. Crucially for subsequent jurisprudence, the Appeals Court was very clear that predatory intent alone is not enough for a liability finding.⁷⁷ It was nevertheless one of the three necessary elements for a successful predatory claim, together with the actual conduct and a ‘dangerous probability’ of successful recoupment. *William Inglis* was also important because it was one of the very first cases, together with *Pierce Packing* and *Janich Bros*, that explicitly referred to a comparison of prices against (variable) costs, in the spirit of the Areeda and Turner (1975) article. Going against that tide though, in *Transamerica* (dealing with IBM’s responses to emerging competition in the ‘plug-compatible’ peripherals market in the early 1980s) the Court of Appeals (Ninth Circuit) specified that there should be no safe legal harbour and that predatory pricing could in principle occur even above average total costs, for instance in the form of limit pricing. The Court believed it ‘should hesitate to create a “free zone” in which monopolists can exploit their power without fear of scrutiny by the law’ (para. 58 of the judgment).

Matsushita is the leading case on collective predation. This case related to the sale of consumer electronics (mostly television sets) in the United States over the 1950s and 1960s.⁷⁸ The allegation was that the 21 Japanese producers engaged in predatory pricing in the United States and cross-subsidised these losses through monopoly pricing in the Japanese market through cartelisation. In spite of a vast amount of documents presented, the District Court could not find sufficient evidence for admissibility of the conspiracy claims. Absent this, predation could only occur as unilateral behaviour, but none of the Japanese producers had sufficient market power to engage in such a practice. The case nevertheless reached the Supreme Court, which, in 1986, stated that predatory pricing conspiracies are speculative by their very nature; and it noted that ‘if the factual context renders respondents’ claim implausible – if the claim is one that simply makes no economic sense – respondents must come forward with more persuasive evidence to support their claim than would otherwise be necessary’ (para. 587 of the judgment). The Supreme Court also stressed the need to show the likelihood of recoupment, else the predatory strategy would be irrational: ‘if predatory pricing conspiracies are generally unlikely to occur, they are

⁷⁷ In line with *William Inglis*, the First Circuit held in *Barry Wright* that predatory intent cannot be a basis for a liability finding.

⁷⁸ This part draws from Elzinga (1989).

especially so where, as here, the prospects of attaining monopoly power seem slight' (para. 590).

Brooke (decided back in 1993) is certainly one of the most prominent judgments in US antitrust case-law on predation. The alleged predator (Brown and Williamson) only had a 12 per cent share in the cigarettes manufacturing market. The allegation related to a fierce price war, which began after the Brooke Group introduced a (commodity-level) unbranded cigarettes pack; as a result, Brown and Williamson introduced its own generic version and started to compete aggressively on price.

The Supreme Court – without embracing the average variable cost rule of *Areeda and Turner* – nevertheless stated that a finding of predatory pricing first required evidence of pricing below some level of cost. Second, and very importantly, recoupment had to be shown to be a 'dangerous probability'. The Supreme Court also stated the principle of recoupment sufficiency: could the alleged predator truly recoup (at least) the entire sacrificed profits in the same relevant market(s) after the predation period? In other words, there should be a quantitative exercise in addition to the qualitative one of determining that the predator has the generic ability to raise the price above competitive levels post-predation.⁷⁹ This imposed a very high standard of proof for the plaintiffs and it was inspired by the Court's concern to avoid type-I errors and the possibility to chill legitimate price competition.

Largely as a result of this new high bar set in *Brooke*, there have been few predation cases in the United States since, and even fewer successful ones for the plaintiffs (Hovenkamp, 2001, counted none at the time he wrote). Bolton et al. (2000) wrote that 'the Supreme Court's *Brooke* decision, [...] as applied by the lower courts, has become virtually a per se rule of non-liability'. Even very recently, Evans (2009) referred to the 'effective elimination of predatory pricing cases in the US'.⁸⁰

One more recent case worth mentioning is *Weyerhaeuser*. Ross-Simmons Hardwood (the plaintiff) and Weyerhaeuser competed in the market for lumber sawmills in the Pacific Northwest region. In 2001, after going into debts of several million dollars, Ross-Simmons closed its

⁷⁹ It is important to note that this case was brought under the Robinson-Patman Act, and not the Sherman Act; however, as noted in Hovenkamp (2005: 346), as the former is 'more expansive' than the latter, any restriction found under the Robinson-Patman Act should equally apply to the Sherman Act.

⁸⁰ In contrast to these authors, Crane (2005) observes that in spite of the high hurdle set in *Brooke*, plaintiffs have kept bringing cases: the author counted at least 57 federal antitrust suits including predatory pricing allegations.

mill. In its antitrust lawsuit, Ross-Simmons claimed that Weyerhaeuser engaged in predatory bidding, that is, it resorted to overpaying for alder sawlogs (an input amounting to about three quarters of total costs) so that prices would rise, as part of a plan to eliminate competition from Ross-Simmons (which could not afford to pay such high prices for its inputs); predatory overbuying was also alleged (that is, Weyerhaeuser would have bought more input than it really required, or about two thirds of the total supplies in the Pacific Northwest). In spite of rising input prices, supply could not expand due to natural resource constraint (annual lumber production was semi-fixed). The District Court ruled in favour of the plaintiff and awarded damages; it also rejected Weyerhaeuser's arguments that the Brooke predatory pricing standards⁸¹ should also apply to predatory bidding. The Court of Appeals (Ninth Circuit) affirmed. But the Supreme Court reversed, on the basis that the Brooke criteria should indeed apply to predatory bidding (*mutatis mutandis*), since the two practices are analytically and practically similar and that the plaintiff itself admitted that it did not meet the *Brooke* standards in its claim.

As Kobayashi (2010) noted, the *Brooke* standards were again tested both in *American Airlines* and in *Spirit Airlines*. In both cases, the courts compared revenues against average variable cost on the routes considered. In addition though, the author remarked, both courts also adopted an 'incremental' version of this test. That is, they 'considered a test that compared whether the incremental profits that resulted from the addition of capacity to certain routes exceeded the incremental costs of adding this capacity'. The courts also opted for opportunity costs, as opposed to accounting cost measures (that is, they considered the revenues foregone by diverting airplanes from more profitable routes to those where predation was alleged). In *American Airlines*, the District Court dismissed the case based, *inter alia*, on the lack of evidence of below-cost pricing (and the Appeals Court affirmed). We consider this case in greater detail in Section 1.6.10.

By contrast, in 2005, in *Spirit Airlines*, the Court of Appeals reversed a District Court judgment, which (by granting summary judgment) had dismissed allegations of predatory behaviour in two leisure passenger

⁸¹ That is, the plaintiff must show competitive injury by proving both that the alleged infringer's prices were below cost and that the defendant had a dangerous probability of recouping its investment in below-cost prices after the prey's exit.

airline markets. The parties subsequently settled.⁸² The allegations against the incumbent airline (Northwest) included both predatory pricing and predatory capacity expansion (adding more flights per day) on the Detroit-Boston and Detroit-Philadelphia routes. The Court of Appeals reversed on the basis that it was wrong to grant summary judgment, given that the evidence produced (when read in the most positive light for the plaintiff, Spirit) called for a full trial. The Court of Appeals was presented with a large amount of evidence. First, Northwest had a high market share on the two relevant routes: prior to Spirit's entry, Northwest had a 72 per cent market share on the Detroit-Philadelphia route (with US Airways being the only competitor), and 89 per cent on the Detroit-Boston and these had been stable over time. Northwest had an incumbency advantage: Detroit-Metro was one of three Northwest's hubs; Northwest controlled 64 out of 86 gates under long-term leases at Detroit-Metro; and it had an overall share of 78 per cent of all outbound passengers. In addition, it appeared that the market exhibited high entry barriers, mostly in the form of scarce access to gates.⁸³ Northwest's price reductions and increases in flight frequency had a significant impact on Spirit's cost structure: Spirit saw its load factor on the Detroit-Philadelphia route drop from 88 per cent in April 1996 (Northwest had not originally reacted to Spirit's entry) to 31 per cent five months later. Being less able to fill airplanes implied a rise in average costs for Spirit, which was in fact forced to leave these routes in September 1996.⁸⁴ Finally, the Appeals Court also noted that '[t]he trier of fact could reasonably find that Northwest recouped any losses from its predatory pricing quickly after Spirit left these routes' (para. 189 of the judgment).

1.5.2 European Union

Moving to Europe, *EC/S/AKZO* is the key case in the European jurisprudence on predatory pricing, dating back to 1985 (we discuss the case in

⁸² See Sagers (2009), who argued that predatory behaviour in the airline industry is far from being an 'impossible' phenomenon and has instead been 'a key tool to preserve market power'.

⁸³ In this sense, for a potential entrant, gates were less of a variable input of production and much more of a fixed barrier (cost) to overcome in order to operate from Detroit-Metro on any given route. Further, it was noted that an entrant would have had to pay 25 per cent more in landing fees than an airline with gates with long-term leases like Northwest.

⁸⁴ As a sidenote, Northwest's lowest unrestricted fares on this route went from \$355 before Spirit's entry to \$49 at the end of June 1996 to \$279 one month after Spirit's exit and up to \$416 in April 1998.

more detail in Section 1.6.9). In *ECS/AKZO*, the European Commission found that AKZO had made a direct threat to ECS (a small competitor), asking it to stop its recent business of selling benzoyl peroxide to the polymer industry, else AKZO would harm ECS through aggressive price cuts in the UK flour additives market, ECS' main business. The threat was then carried out. The European Commission explicitly rejected the Areeda-Turner test which we discussed above. Price-cost analysis was only one of the elements to be assessed, it argued; valuable information on the potentially abusive conduct comes instead by contextualising the pricing practice within a wider exclusionary strategy, so that in principle even prices above average total costs could be predatory. On appeal, the Court of Justice upheld the Commission's decision, but it also made further clarifications: first, prices below average variable cost must be regarded as abusive; second, a price between average variable cost and average total cost would be deemed predatory only if this was part of a wider strategy aimed at eliminating a competitor, that is, if there was evidence of intent. This judgment effectively set the European standard on predation.

Another European landmark case on predation is probably , which involved a number of exclusionary practices identified by the Commission. Insofar as predatory pricing in the market for non-aseptic cartons was concerned, the Commission found an infringement with respect to the Italian market, where sales at a loss were judged to be part of a deliberate strategy aimed at eliminating competitors. The courts agreed. The Court of Justice further clarified that there was no requirement (at least in that case) for the Commission to prove the likelihood of recoupment, that is, that the alleged predator would have the ability to offset its short-term losses (arisen during the predatory process) through subsequent profits once the prey had been eliminated from the market.

Contrary to the US though, European jurisprudence on predatory pricing did not become scantier from the early 1990s. Indeed, we have observed several recent cases on predatory pricing decided either by the European Commission or by the national authorities, some of which made their way up the courts.⁸⁵ One reason for this jurisdictional divergence

⁸⁵ In *Wanadoo*, the European Commission found that France Télécom (through a subsidiary, Wanadoo Interactive) had engaged in predatory pricing in high-speed internet access services to residential users in the early 2000s. In its investigation, the Commission applied a price-cost test. It also examined evidence of intent, recoupment (although it clearly stated that this was not a necessary condition for a finding of predation) and effects of Wanadoo's conduct on competitors. The Commission's decision was upheld by both the General Court, which *inter alia* noted that a finding of predation does not require the

is that in Europe, as stressed in *Tetra Pak II*, there is no need to prove the likelihood of recoupment, considerably lowering the standard of proof vis-à-vis the US. Another reason can be found in the wave of liberalisation that took place in Europe in the 1980s and 1990s. Suddenly, markets that were previously state monopolies (such as railways, postal sector, telecommunications, energy) opened up to competition, but players were clearly very asymmetric, with a strong incumbency advantage and persistent dominant positions enjoyed by the former monopolistic firms, which at times resorted to pricing and contractual practices to avoid competition from new entrants. Several of these exclusionary cases will be reviewed in this book, starting with some recent predation cases in the next section.

1.6 Cases

In this section we review a number of international cases. We shall focus in particular on those where scale economies and strong initial advantages on the side of the incumbent seem to play an important role, in a way that is consistent with the theoretical considerations we made earlier in this chapter. Of course, highlighting industry elements that broadly match the conditions of economic models of predation is not enough to claim that there has been actual predation; they simply suggest that there may be a theory of harm which supports the allegation of predation. To make a precise statement on the merit of a case, we should have access to more detailed information (especially on relevant cost measures) that is usually not publicly available. Likewise, as mentioned in Section 1.5.2, EU competition law does not require proof that there is a high probability of recoupment, implying that the available documents do not shed much light on this issue. Still, in what follows we shall try to discuss to what extent recoupment may have been possible in some of the EU cases.

In any event, the message of this chapter is not that predation is a pervasive phenomenon that arises whenever there exist scale economies

effects of the predatory conduct to be demonstrated) and by the Court of Justice. We review this case in more detail in Section 1.6.5, where we also discuss the application of the price-cost test applied in this case.

At the time of writing, the European Commission is investigating a number of potentially anti-competitive practices by Qualcomm. In December 2015, the Commission announced that it had taken the preliminary view that, *inter alia*, between 2009 and 2011 Qualcomm engaged in predatory pricing by selling certain baseband chipsets at prices below costs, with the intention of hindering competition in the market (see European Commission, press release IP/15/6271 of 8 December 2015 and Case 39.711 *Qualcomm*).

and an incumbency advantage. Rather, we are interested in identifying those (likely exceptional) situations where predation is likely to occur, as well as those where predation will be unlikely.

1.6.1 Napp

We begin our review with a landmark case from the UK, *Napp*. This is a case which can be interpreted in the spirit of the two-sided markets version of the model we presented in Section 1.3.2.8 (see Section 1.2.3.2 for a non-technical discussion of this model). In 2001, the UK Office of Fair Trading ('OFT') found that Napp, a pharmaceutical company, had abused its dominant position in the market for the supply and distribution of sustained release morphine in the UK. This infringement involved both a charge of predatory pricing in the hospital segment and one of excessive pricing in the community segment (we explain the nature of these two groups of buyers below).

Napp had an incumbency advantage at the beginning of the predatory period, with a market share of 80 per cent in the hospital segment and of over 90 per cent in the community segment. This original advantage arose from patent protection over the period from 1980 to 1992. This was reinforced by moderate entry barriers upon patent expiry which took the following forms: the time needed to acquire the authorisation to manufacture the drug; a separate authorisation to market the drug; and the need to build a reputation as a trusted brand for a given drug. Marketing campaigns were considered to be costly: direct marketing (also known as detailing) to general practitioners ('GPs', or family doctors) was a very expensive endeavour given how numerous and scattered they were.

While it may appear odd that Napp could engage in too low prices in a market segment and too high prices in another market segment, our theory helps interpret the case, as we next discuss.

Sustained release morphine was sold to two completely different groups of buyers. One group was represented by hospitals, which were very price-sensitive (pharmaceuticals had to be paid out of their budget) and could count on the advice of specialist doctors for an assessment of quality of the competing products. The other group (about eight times as large, by volume sold) was represented by the community segment, where buyers were GPs who prescribed products for their patients (with the National Health Service paying the bills, thus making these buyers not particularly price-sensitive). Moreover, GPs – not being experts (as well as to facilitate

treatment continuity) – tended to select those products which had already been chosen by hospitals (further raising entry barriers in the segment where price competition was less fierce).

The facts in *Napp* can therefore be consistent with an asymmetric two-sided market, where hospitals mostly care about prices (and do not care about choices made by GPs), while the demand of the community segment strongly depends on the choices made by hospitals. As discussed in Section 1.3.2.8, an incumbent like Napp – who has already acquired a solid reputation – may want to sell below costs to the crucial side of the market, where entry barriers are lower (the hospital market in this case) to make sure the rival does not gain a significant foothold there. As a result, it would deter the rival's activity also on the other side of the market (in this case, the community segment, where entry barriers were much higher), whose demand closely followed the choice made by hospitals. Thus, the incumbent can behave like a monopolist on the community side of the market, recouping any losses made to win the other side (the hospitals' segment).

Indeed, Napp's pricing strategy had an appreciable effect on its main competitors. While their combined market share in the community segment hovered around 4 per cent over the predatory period, the product of Napp's only significant competitor (Boehringer Ingelheim), Oramorph, saw its market share in the hospital segment decline from 20.1 per cent in 1997 to 4.1 per cent in 2000.

One of Napp's key defences for its low prices in the hospital segment was that there was a 'follow-on' effect that had to be taken into account: it estimated that one unit sold to the hospital segment led on average to 1.35 units sold in the community segment. The OFT remained unconvinced for a number of reasons (see pages 50–4 of the decision). The main one was that Napp's aggressive discounts in the hospital segment were only offered on drug strengths where there were competing alternatives and in geographic areas where there was a competitor. But if the follow-on effect had been the key motive for Napp's pricing policy (given that, in principle, even a monopolist would offer a low price on one side of the market in order to attract both sides), there should have been aggressive discounting across all product categories and geographic areas.⁸⁶

⁸⁶ The UK Competition Appeal Tribunal also pointed out that there was no contemporaneous evidence (internal documents) suggesting that Napp adopted this pricing strategy to benefit from the claimed 'follow-on' effect; rather, this was put forward *ex post* during the proceedings.

1.6.2 Sanofi-Aventis

Sanofi-Aventis, a French case, had several features in common with *Napp*. Yet, there were some critical differences which (correctly, in our view) led to a different outcome in that case.

Sanofi-Aventis was found to have a dominant position in the market for low molecular weight heparins ('LMWH', an anti-coagulant). Since 2001, it had been offering Lovenox (its branded LMWH) to hospitals free of charge. The theory of harm put forward by the French Economy Ministry before the French Competition Authority was similar to that in *Napp* (the case we have just discussed above): Sanofi-Aventis was allegedly engaging in predatory pricing in the hospital segment so as to be able to lock patients in for treatment once out of the hospital (through a 'sourcing effect'); and to subsequently charge supra-competitive prices in the community segment. As evidence, the Ministry showed an increase in Lovenox's market share (by volume) in the hospitals' segment (for prevention use): from 35–40 per cent in 1998 to 65–70 per cent in 2004. The Ministry also pointed towards an increase in Lovenox's market share in the community segment.

This is where the high-level similarities with *Napp* ended. In fact, there were three crucial differences.

First, Sanofi-Aventis faced competition by two large, well-established global players (including Pfizer, with a 15–20 per cent share of the community segment by the late 1990s) and by a mid-sized player, Léo Pharma (with a 5–10 per cent share). This raises doubts about the rationality of a predatory strategy.

Second, it is difficult to see an intention to pursue a predatory strategy, since Sanofi-Aventis did not start, but simply followed the rivals in their price reductions in the hospital segment. As a further caveat, the increase in Sanofi-Aventis' market share in the community segment between 2001 and 2006 actually reflected the merger between Sanofi and Aventis (the sum of the separate market shares was even slightly larger than that of the merged entity).

Third, the link between the hospital and the community segment was weak⁸⁷ and definitely less clear than in *Napp*. The key fact which was

⁸⁷ This was probably due to the fact that while there was no price regulation in the hospital segment, prices in the community segment were negotiated between the pharma companies and a specialised public authority (CEPS). The French Competition Authority noted that drugs that entered the market later commanded a lower price, possibly due to more uncertainty over their effectiveness. By contrast, hospitals would simply choose the cheaper LMWH drug available.

inconsistent with the theory of harm proposed (via the ‘sourcing effect’) was that the collective market share (by volume) in the community segment of the large firms offering hospitals LMWH drugs free of charge actually *dropped* by about 20 percentage points (85–90 per cent to 65–70 per cent) between 2001 and 2006. (That is, contrary to the rationale behind the proposed theory of harm, heavy discounting in the hospitals’ segment by some companies did *not* result in a higher market share in the community segment.) A large fraction of this drop was taken up by Léo Pharma: its share of the community segment grew from 10–15 per cent to 20–25 per cent over the same period. Léo Pharma achieved this in spite of no longer offering its LMWH drug free of charge to hospitals and in spite of its share of the hospital segment being below 5 per cent for the entire period. This element cast a serious doubt over the supposed strength of the ‘sourcing effect’.

Eventually, the French Competition Authority dismissed the case and – based on the public information available – this seems a reasonable decision to us.

1.6.3 Cardiff Bus

In November 2008 the UK Office of Fair Trading (‘OFT’) found that Cardiff Bus had infringed Chapter II of the UK Competition Act by engaging in predatory conduct in Cardiff’s local bus market. In response to 2 Travel’s entry into the market with a new no-frills bus service, Cardiff Bus introduced its own no-frills bus service (the ‘white service’), running on the same routes and at similar times of the day as 2 Travel’s services (with the added value of a single ticket valid on the whole Cardiff Bus network at no extra charge). The OFT showed that the white services were run at a loss (their revenues were not even sufficient to pay the wages of the drivers on those routes) until shortly after 2 Travel’s exit, when Cardiff Bus discontinued them altogether (and raised fares on other services). 2 Travel was facing difficult financial conditions, up to the point where it failed to comply with safety requirements for lack of funds available (most of its buses had been bought second-hand); this contributed to its early exit.

The economics behind this case could be read in the spirit of the predation mechanism we proposed in Section 1.3. First, there appeared to be significant economies of scale and scope, partly driven by network effects: profitable entry was only attainable at a certain scale, that is, only once a critical network size was in place (as consumers value the combination of schedules and routes). The OFT noted that scale

economies at a network level were compounded by the significance of scale (density) economies at the level of a single route, as consumers value frequency of services. Second, barriers to entry were also found to be high, due for example to limited bus depot availability. Reflecting such entry barriers, 2 Travel was actually the first entrant since 2000.

Cardiff Bus' incumbency advantage took the form of its established routes' network. In the Cardiff area, it controlled about two thirds of total network traffic, with the competition being geographically very fragmented, and no other single operator having a sizeable share of the network; moreover, it had 72 per cent of the services (by frequency) stopping at Cardiff Central Bus Station. Drawing from a report it commissioned, the OFT wrote that:

[O]nly via large scale, multi-market entry, might an entrant deprive the incumbent of its multi-market advantage, and thus avoid being selectively targeted. (Para. 6.47)

Against this structural setup, 2 Travel's entry with only four routes was fairly limited (Cardiff Bus had 42 different services); it had 20 buses operating within a 30-minute journey from Cardiff Central Bus Station (Cardiff Bus had 223). Speculating, one could argue that this was the first stage of a wider entry strategy, but Cardiff Bus' predatory strategy choked such staged entry while still nascent.

In principle, one could also read this case in the spirit of a reputation model: the OFT, having collected testimonies from competitors, also drew the conclusion that Cardiff Bus had built a reputation (mostly among smaller operators) for fiercely fighting entry or expansion, especially on Cardiff city routes.

The OFT discussed recoupment in this case and found that recoupment was very likely, due *inter alia* to a combination of high barriers to entry and Cardiff Bus' reputation.⁸⁸

1.6.4 Telecom Italia

Another interesting European case brought by a national competition authority is *Telecom Italia – Comportamenti abusivi*. In 2004 the Italian Competition Authority (the 'Authority') found that Telecom Italia ('TI'), the public monopolist before the liberalisation process, had abused its

⁸⁸ See chapters 6 and 7 of the OFT decision.

dominant position by using a variety of practices (in what follows, we focus on the price-related abuses). The case was concerned with the provision of services to government bodies (typically tendered through the CONSIP agency) and large business customers.

From a legal perspective, the Authority did not challenge TI's retail prices as predatory (see para. 262 of the decision). By contrast, it sanctioned TI's retail offers to the above customer categories that could not be matched by TI's competitors. These were companies that were themselves customers of TI, buying wholesale services at regulated prices. The Authority stated that these abusive retail offers could also be interpreted as illegal price discrimination at the wholesale level (in that TI was implicitly 'charging' its downstream affiliates a discriminatory price with respect to the price offered to the other telecom operators).

From an economic perspective, however, the case may be read in the spirit of the model we presented in Section 1.3.1 (see Section 1.2.3.1 for a non-technical discussion).

Evidence referred to during proceedings suggested that TI set prices in a selective and aggressive way, seeking to take away key customers (the so-called 'key contributors') from competitors, thereby hindering rivals' expansion.⁸⁹ The Authority also pointed to the existence of significant scale economies enjoyed by TI, strengthening its incumbency advantage over new entrants, which still had to build up or fully develop their infrastructure (viable only if they reached sufficient scale) and customer bases.⁹⁰

Further, the presence of staggered tenders by government bodies and large business customers likely meant that buyers were not able to coordinate or pool their purchases. In Section 1.2.3.1 (and technical Section 1.3.2), we explained how such a market feature (customers buying at different times) may make exclusion potentially easier to implement, all else equal.

1.6.5 Wanadoo

Another interesting case from the telecom industry is *Wanadoo* (known on appeal as *France Télécom*). In 2003, the European Commission found that

⁸⁹ Internal documents showed TI's management was willing to incur losses in order to win – or win back – important business customers (see para. 122 of the decision).

⁹⁰ For instance, at para. 275 of the decision, a cable rival, Fastweb, argues that Telecom Italia's strategy aimed at eliminating competitors' incentives to invest in new and non-recoverable alternative telecom infrastructure, with the ultimate effect of inhibiting the development of competitors in the long-run.

France Télécom (through a subsidiary, Wanadoo Interactive, henceforth 'Wanadoo') had engaged in predatory pricing in high-speed (ADSL) internet access services to residential users.

In the early 2000s, the broadband market in France (as elsewhere in Europe) was at an infant stage. Operators had to overcome consumers' doubts over a new technology and establish a reputation as providers of high-quality and reliable services (in addition to having to offer value for money, given the presence of narrowband, or dial-up, alternatives). Operators who were most successful at this would gain an early foothold in the market and thus a competitive advantage. As the Commission put it:

[...] Service providers must, during this high-speed market development phase, build an image as the default supplier of a product viewed by the consumer as technically sophisticated and become large enough to benefit from economies of scale.

In this process, the chronological sequence of entry into the market is far from neutral. Clearly, a service provider that has a considerable head start over its competitors during the initial phase of market growth is able to capitalise on the momentum thus gained. By contrast, laggards must make a much bigger effort to acquire customers if they wish to make up for lost time and bridge the resulting image gap and confer on their high-speed service the same notoriety as that of the dominant undertaking's flagship offering. In these circumstances, new competitors are confronted with the need not only to carry out the expenditure technically necessary in order to provide the service but also to undertake substantial advertising and promotional expenditure both to raise their product's profile and to undermine loyalty to the dominant undertaking's brand. (Paras 351–52 of the decision)

In the context of nascent technological markets, it is not uncommon to observe prices that are *prima facie* 'low' (this practice is often referred to as penetration pricing). These markets often involve consumer goods where end-users also need to purchase some expensive hardware to benefit from a service (for example, a player, a games console, a TV decoder or set-top box, an internet modem). Penetration pricing can therefore incentivise consumers to buy earlier on when there is uncertainty about the true value of the consumer good or service. Operators, from their side, benefit from early uptake also through scale and network economies.

Against this background, the Commission first carried out a price-cost test in *Wanadoo*. It showed that retail prices were below average variable costs between March 2001 (when the mass marketing of the ADSL services began) and August 2001; and that in the following period up to October 2002 they were approximately equivalent to variable costs, but significantly below total costs.

A first difficulty with this case was the determination of the cost items to include in the relevant cost benchmark (variable cost). Wanadoo claimed that its marketing costs had a long-term nature and were fixed. The Commission disagreed (see paras 61–6). In particular, it stressed that even abstracting from general marketing, Wanadoo incurred ‘advertising and marketing [costs] specific to Wanadoo Interactive’s ADSL services’ and that these were ‘quite specific and correspond[ed] to successive campaigns which focus[ed] strongly on sales promotion’; the Commission also pointed towards a strong correlation between advertising campaigns and new subscriptions rates. Further, as part of the customer acquisition costs, it highlighted the existence of ‘special offers’, which ‘by their very nature [...] are directly linked to the purchase of each new product unit sold (discounts, refund offers, etc.)’. (There was no mention, however, of direct subsidies to customers in the form of hardware such as modems, within the Commission’s discussion of marketing costs.) In sum, the Commission thought that these costs should be treated as variable.

A second, related, problem was the number of years over which such costs should be depreciated (the Commission opted for 48 months, that is, the average length of customer tenure, see paras 76–9).⁹¹

Next, going beyond a pure price-cost analysis, the Commission (and then the General Court) went three steps further. First, the Commission sought to establish the intent of the predatory strategy.⁹² To do so, it relied on the internal company documents it discovered. Most of the relevant quotations were omitted in the Commission’s decision for confidentiality reasons (paras 110–17). Some, however, can be found in the General Court’s judgment, which itself refers to documents discovered by the Commission. These include:

- the framework letter for 2001 containing the following wording: ‘our pre-emption of the ADSL market is imperative’; [and]
- a note of [Wanadoo]’s strategic management of December 2001 stating: ‘The high-speed and ADSL market will, for the next few years, continue to be conquest-driven, the strategic objective being to gain a dominant position in terms of market share, the period of profitability only coming later.’ (Paras 199 and 215 of the General Court’s judgment)

⁹¹ The Commission, however, noted that for the period considered in the decision, ‘the length of the period over which the non-recurrent costs are to be spread is immaterial to the finding that the full costs are not covered, because the full costs are not covered even if depreciation takes place over an infinitely long time’ (para. 79).

⁹² The Commission needed to establish the intent of the predatory strategy only for the period during which the prices of Wanadoo were above AVC but below ATC.

As with many quotations taken from internal business documents, it is difficult to understand the full context and how they should be interpreted through an antitrust lens.

Second, differently from US law, to establish predation, there is no requirement in the EU to show that a dominant firm engaging in below-cost pricing can actually recoup the lost profits once the prey has left the market (see the Court of Justice's judgment on this very case). The Commission nonetheless looked at this aspect and found that recoupment was plausible, because of high entry barriers (paras 332–67), such as: disincentives for existing customers to change supplier; the costs for a potential competitor to enter and acquire critical size; and the major difficulty in replicating a telecommunications network and entering upstream. The Commission further pointed to Wanadoo's margin levels post-October 2002 as part of the evidence making recoupment a plausible argument.

Third, there was a brief analysis of the effects of Wanadoo's pricing behaviour on competitors. The General Court (paras 259–67) noted that Wanadoo's market share (in residential access to broadband internet) increased from 50 per cent (in March 2001) to 72 per cent (in August 2002), before falling to 63.6 per cent (in October 2002). The nearest competitor had an 8 per cent share, while all others had less than 3 per cent. Mangoosta, another competitor, left the market. Cable operators also saw their market share decline significantly. Further, the General Court added that there had been few and insignificant instances of entry, which the Commission and the General Court attributed to Wanadoo's aggressive pricing. Some operators, however, saw their market share slightly increase during the predatory period (see para. 385 of the Commission's decision). The Commission dismissed this fact on the basis that their share could have grown more absent the abuse (see Chapter 2 – Section 2.6 – for a discussion of similar 'counterfactuals' in competition law enforcement).

It is not clear to us, from the decision and the judgments, whether Wanadoo's behaviour was predatory. The above facts (both the market features and the Commission's findings) are, however, compatible with a framework other than a predatory one: competition *for* the market. As stated above, the broadband market in the early 2000s was clearly a nascent, growing market, with strong network externalities. Operators getting an early foothold in this market would be the likely 'winners', that is, obtain a strong competitive advantage and thus quickly gain market share.

Network externalities are indeed a key ingredient of the model described earlier in this chapter (Section 1.3.2.7). On balance, though, it is unlikely

that Wanadoo's strategy was exclusionary with reference to the models presented earlier in this chapter. We saw in Sections 1.2.3.2 and 1.3.2.3 how in the presence of growing markets, predation is a *less* likely equilibrium outcome, as the large 'prize' in future periods can be sufficient to stimulate entry: in 2001, expectations about growth in broadband uptake were very high (growth which indeed materialised over the years to come).

Moreover, as stated at the outset, low prices at an early stage can help consumers overcome uncertainties about the adoption of a new technology and there are efficiencies associated with such pricing strategies.

In sum, it is not clear whether the facts in *Wanadoo* should have led a predatory finding; and it is even less clear which coherent theory of harm could support it. More generally, an aggressive antitrust stance in a nascent, fast-developing market could stifle incentives to innovate in the first place and thus reduce consumer welfare over the long run; as such, caution would therefore be required.

1.6.6 Deutsche Post

Shortly before, in 2001, the European Commission found that Deutsche Post AG ('DPAG') had abused a dominant position in the market for mail order parcel services. The Commission argued that by making use of predatory pricing and fidelity rebates, DPAG tried to prevent competitors in the mail-order service from developing the infrastructure needed to compete successfully.

The notion of costs used by the Commission in the decision was that of 'average incremental costs' (or 'AIC').⁹³ This explicitly excluded all common costs that would be incurred upon the provision of other services. DPAG's network included freight centres and delivery points. This infrastructure was used for a variety of services. The cost of capital and maintenance for common sorting infrastructure could not be attributed to any business line in particular and was therefore excluded from the AIC of mail-order sales. The Commission considered several

⁹³ This was in line with the European Commission's Notice on the Application of the Competition Rules to Access Agreements in the Telecommunications Sector, where the Commission suggested that the examination of average incremental costs (over a longer period than one year) could be more suitable for network industries, since these industries have much larger common and joint costs compared to most other industries (see OJ [1998] C 265/02, paras 113–15).

cost categories. The next excerpt (referring to the sorting costs⁹⁴) highlights the difference between service-specific and common costs quite clearly:

The capital costs of setting up the 33 freight centres and 476 delivery points cannot be attributed to a particular service. These costs will be incurred as long as the statutory obligation to meet demand to legally required service-quality standards applies. The staffing and equipment costs of sorting, on the other hand, are entirely dependent on the actual volume of parcels to be conveyed. Thus the staffing and equipment costs of activities dependent on the volume processed can be attributed in direct proportion to the mail order parcel service. (Para. 13 of the Commission's decision)

Another key dimension of this cost benchmark was the temporal horizon over which to consider a cost to be incremental. The Commission considered the six-year period from 1990 to 1995, that is, coinciding with the alleged predatory period. For a more thorough discussion of the case, and of the Commission's use of AIC, we direct the reader to OFT (2005). The following quote therein contained suggests the rationale of adopting such a cost measure and such a temporal horizon:

Over six years all overheads (ie: costs that on a short term basis do not vary directly with output) were deemed avoidable because this was ample time to realize that the mail order parcel service was not covering such costs and therefore raise prices or dispose of the additional assets required to provide that service. To put it another way, over this period DPAG would have had the choice not to re-incur all incremental costs of running its mail order parcel services business. (OFT (2005), para. A.21)

Having thus defined AIC, the Commission found that DPAG's prices in mail order parcel services were below such cost measure in each year between 1990 and 1995 and ruled against Deutsche Post.⁹⁵

In light of the above considerations, the AIC defined in *Deutsche Post* seems to us, *mutatis mutandis*, more akin to a good proxy for average

⁹⁴ The other service-specific costs included by the Commission were: collection, regional and local transport (proportionally to mail order parcels volumes); delivery (specifically, the handing over of a mail-order parcel), as opposed to driving, which was considered a cost that would be incurred as part of its statutory requirements in relation to letter deliveries.

⁹⁵ Eventually, the Commission only imposed a fine for DPAG's rebates scheme, but not for its predatory pricing (since it was the first time the Commission applied the notion of AIC in an abuse case). Moreover, DPAG offered commitments that amounted to a structural separation of its commercial parcel services from its letter business.

variable cost (for the case of a firm producing multiple products) than for average total cost.⁹⁶

Therefore, according to our proposed policy rule (see Section 1.4), there should have been a strong presumption of predation at the outset of *Deutsche Post*, purely based on a strong dominant position and on prices being below variable costs (profit sacrifice). In addition, though, we believe that it is useful to attempt to interpret the facts of that case in the spirit of the predatory mechanism we proposed in Section 1.3.2.6, which may contribute towards a coherent theory of harm.

The scale and scope economies in the postal sector were significant, and DPAG had an incumbency advantage: DPAG was the former state monopolist and, as such, could rely on a fully developed distribution infrastructure and on exclusive right in the market for letters and small parcels. Further, DPAG had a stable market share in excess of 85 per cent in the market for mail-order parcels for the whole of the 1990s. Also, the following quote from the Commission's decision suggests that the incumbent's pricing policy had the ability to deprive rivals of the scale and scope economies required to compete effectively:

Contrary to what DPAG maintains, all of the disputed fidelity rebates are likely to have an effect on the opportunities that other suppliers of mail-order parcel services have to compete. Successful entry into the mail-order parcel services market requires a certain critical mass of activity (some 100 million parcels or catalogues) and hence the parcel volumes of at least two cooperation partners [customers with very large orders] in this field. By granting fidelity rebates to its biggest partners, DPAG has deliberately prevented competitors from reaching the 'critical mass' of some 100 million in annual turnover. This fidelity rebating policy was, in precisely the period in which DPAG failed to cover its service-specific additional costs (1990 to 1995), a decisive factor in ensuring that the 'tying effect' of the fidelity rebates for mail-order parcel services maintained an inefficient supply structure [...]. (Para. 37 of the Commission's decision)

The rationale of DPAG's strategy might therefore have been to prevent competitors from achieving the efficient scale they needed to expand in the market of mail-order parcel services.⁹⁷

Moreover, given the existence of important common costs with other postal services, mail-order operators could later start to compete with other

⁹⁶ We direct the interested reader to Sappington and Sidak (2003) for an insightful piece, which includes a lengthy discussion of the *Deutsche Post* case.

⁹⁷ While the above excerpt refers to fidelity rebates (and not predatory pricing), the economic mechanism at work is the same (i.e. making the entrant unable to reach a critical mass), as we shall argue in Chapter 2.

services offered by DPAG. Hence, as informally discussed in Section 1.2.3.2 (and formally shown in Section 1.3.2.6), DPAG might have sought to price aggressively in the ‘early’ markets to preserve a dominant position in all the markets where it operated. Indeed, after the end of the abuse found, some competitors managed to reach such critical mass and operate across adjacent markets. For instance, Hermes Versand Service was initially created for the mail-order trade’s own use, but its infrastructure was later used to convey parcels for third parties and by 2000 it became the sixth largest courier, express mail and parcels operator in Germany (see para. 38).

1.6.7 Aberdeen Journals

In *Aberdeen Journals*, the UK Office of Fair Trading (‘OFT’) found that Aberdeen Journals adopted predatory behaviour in March 2000 in the pricing of advertising space in its local newspapers, in an attempt to drive the Aberdeen & District Independent newspaper (the only competitor) out of the market.⁹⁸ The OFT could not condemn the abusive practice over an earlier period as the Competition Act 1998 only entered into force on 1 March 2000. So, it is not that the predatory practice necessarily lasted only for one month, it is that only a certain period out of the all duration of the practice could be sanctioned. The Competition Appeal Tribunal (‘CAT’) upheld.⁹⁹

The relevant market was that of the supply of advertising space in local newspapers (both paid-for and free) within the Aberdeen area. Aberdeen Journals was found dominant, with a market share in excess of 70 per cent (by value).

The OFT found that the revenues received from advertisers by Aberdeen Journals for its Herald & Post did not cover its average variable costs of production (this weekly was offered free of charge to readers). The pagination (that is, the number of pages) of the Herald & Post increased significantly after the launch of the Independent, from under 100 pages per

⁹⁸ A previous decision by the OFT in 2001 had been set aside by Competition Appeal Tribunal due to an unsatisfactory definition of the relevant market.

⁹⁹ Whish and Bailey (2015: 794–5) suggested that the CAT judgment contains various elements of special interest, including: the cost-based rules set in *AKZO* and *Tetra Pak II* should not be applied ‘mechanistically’; the timeframe over which costs are calculated is important, as the longer this is, the likelier the chance that costs should be assessed as variable and not fixed; and there could be, exceptionally, an objective justification for below-cost pricing by a dominant firm, but this rebuttal is unlikely to be accepted if the pricing in question is part of a wider exclusionary strategy or follows an aggressive entry by a competitor.

month in the period from October 1995 to April 1996 to 148 pages in May 1996. Pagination then peaked between September 1998 and October 1999, at around 350–480 pages per month.

As for barriers to entry and expansion, the OFT – referring to earlier analyses of the newspaper industry by the UK Competition Commission – suggests that they are significant:

The [UK Competition Commission] has found that the initial cost of launching a free newspaper is relatively low, with desktop publishing technology becoming readily available, but has recognised that the cost of sustaining or expanding a presence can be higher (particularly where there are calls for an expansion of the editorial content). In particular, a new entrant might expect to incur losses during its first years, the time taken to establish credibility and acceptability with advertisers (the only source of revenue for free titles). A free paper must establish credibility with readers to be able, in turn, to attract advertisers. Further, establishing distribution networks can be difficult and time-consuming. (Para. 145 of the OFT's decision, footnotes omitted)

Further, based on the information available to us, Aberdeen Journal had an incumbency advantage, consisting essentially of an established distribution network, a set of local newspapers and journalists and a sound reputation with advertisers. The entrant clearly did not.

One may therefore argue that by preventing the Independent from obtaining revenues from advertisers at an early stage, Aberdeen Journal could limit its ability to expand (or even be viable) at a later stage. This effect was likely to be compounded by a feedback mechanism via readership: lack of the rival newspaper's credibility also implied fewer readers and thus fewer advertisers, in turn limiting revenues.

These ingredients are broadly consistent with those we highlighted in Section 1.3.2.1, where we discussed how an incumbent could prevent an existing rival or a new entrant from expanding, in the presence of scale economies.

In addition, the OFT referred to an earlier remark made by the UK Competition Commission on the newspaper industry suggesting that reputation effects might also be at work:

The [UK Competition Commission] has also recognised that the behaviour and reputation of an incumbent influences ease of entry. In particular, it notes that established newspaper publishers can respond to new entrants in several ways and that the expectation of the incumbent's response can deter potential entrants and, if entry does take place, could reduce the entrant's chances of success. [...] A reputation for predatory response to new entry creates a barrier to entry. This is particularly important where the incumbent operates in several product or

geographic markets: [...] Northcliffe [Aberdeen Journal's parent company] is active in many geographic markets. Accordingly, in this case, the predatory response of Aberdeen Journals to entry could create a reputation that acts as a significant barrier to entry and maintains its position in this and other markets. (Paras 146–7, footnotes omitted)

Overall, though, the robustness of a theory of harm in this case exclusively based on reputation effects is unclear to us, based on the information available.

1.6.8 Media24

Another interesting case related to local newspapers is *Media24*, which provides a useful basis to discuss the relevance of cost benchmarks and of any evidence of predatory intent.¹⁰⁰

Following a reference by the Competition Commission of South Africa ('the Commission'), the Competition Tribunal of South Africa ('the Tribunal') found that Media24, a media group, used its Forum community newspaper anti-competitively, driving the Gold Net News ('GNN') community newspaper out of the local market in the town of Welcom.

According to the account in the Tribunal's judgment, at the time of the alleged predatory conduct, there were three community newspapers in the town of Welcom. Media24 owned two of these, Forum and Vista, with a combined share in the relevant market of about 75 per cent. GNN, an independent competitor, had the remaining share of the market. These community newspapers were free of charge to readers and raised revenues from advertising.

The Commission alleged that Media24 used Forum as a 'fighting brand', by selling adverts below cost and with the objective of driving GNN out of business. While the allegation of predatory conduct related to the period from 2004 to 2009, the Commission also pointed to evidence of Forum being loss-making from 2001 (shortly after GNN's entry) until 2009, when GNN exited the market. Media24 then closed Forum in January 2010.

One important aspect of this case is what was referred to in the proceedings as the 'cannibalisation effect': by keeping Forum's prices for adverts 'low', Media24 was forgoing some advertising revenues that Vista would have earned, as some of the advertisers in Forum would have switched to Vista (which charged higher prices), had Forum's prices been

¹⁰⁰ At the time of writing, this judgment is under appeal at the Competition Appeal Court of South Africa.

higher. The Tribunal noted that these foregone revenues represented an economic cost to Media24 and were therefore considered in detail in the Tribunal's assessment, as we shall see below.

The Tribunal was asked to analyse this case under two provisions of the Competition Act of South Africa (the 'Act'): Section 8d(iv), which prohibits dominant firms from 'selling goods or services below their marginal or average variable cost' (unless the firm concerned can show technological, efficiency or other pro-competitive gains which outweigh any anti-competitive effects); and Section 8c, which prohibits dominant firms from engaging in an 'exclusionary act' (if any anti-competitive effect outweighs any technological, efficiency or other pro-competitive gain).

This distinction is insightful in our view, both in policy terms and for the assessment of this case, and we shall return to this below.

Pricing below average variable cost or marginal cost The Tribunal noted that the economists on both parties (Commission and Media24) agreed that the appropriate cost benchmark to consider from an economic perspective, given the facts of the case, was that of average avoidable cost ('AAC').

There was substantial disagreement between the parties on the level of Forum's AACs. The Tribunal found that whether Forum's prices exceeded AACs depended on the assumptions on the cannibalisation effect, which we introduced above. But due to the significant uncertainty on its magnitude, the Tribunal found that the Commission had not established on a balance of probabilities that Forum's revenues were below its AACs during the period of the alleged predation.

In sum, the Tribunal dismissed the Commission's allegations around an infringement of Section 8d(iv) of the Act as there was insufficient evidence that Forum's prices were below the relevant cost benchmark.

Exclusionary conduct To reach a view as to whether Forum's conduct was exclusionary under the meaning of Section 8c of the Act, the Tribunal considered various categories of evidence. Next, we summarise the Tribunal's findings on: (i) the relationship between Forum's prices and its average total costs; (ii) direct predatory intent; and (iii) indirect predatory intent.¹⁰¹

¹⁰¹ The Tribunal also considered the issues of recoupment and of whether GNN was an equally efficient competitor. The Tribunal noted that evidence of recoupment is not required by the Act, but in the context of Section 8c of the Act this evidence 'can either serve to bolster a conclusion reached around intent or to serve to throw some doubt on its reliability' (para. 491). The Tribunal, inter alia, pointed to evidence adduced by the

First, on the relationship between Forum's prices and its average total costs, the Tribunal found evidence that Forum's prices failed to cover average total costs over the whole period of the alleged predatory conduct.

Second, in relation to evidence of direct predatory intent, the Tribunal reviewed internal documents dating back to the period of the allegations and heard various witnesses. The Tribunal pointed to a Forum manager and his superior having 'many discussions [...] about using Forum as a vehicle to prevent GNN expanding in the market. These discussions were never documented to avoid them falling into the hands of the Commission' (para. 290). The Tribunal also noted that the same Forum manager 'emphasised that since he knew his own titles costs, and that Forum was not profitable, he confidently assumed that GNN, lacking Media 24's economies of scale and scope could not be making money. Indeed he was surprised the paper lasted as long as it did' (para. 297 of the Tribunal's decision). He also stated, the Tribunal reported, that 'Forum was kept alive [...] to keep GNN out of the market, if it hadn't been there [...] Media 24 would have closed it' (para. 298).

The Tribunal concluded that 'the evidence of direct intention to use Forum as a predatory vehicle against GNN [was] supported, not only by the oral testimony of an erstwhile employee, but also documentary evidence in the record throughout the period' (para. 372).¹⁰²

Third, the Tribunal considered evidence of indirect predatory intent. It considered three categories of such evidence, and found the following:

Commission showing that (i) Vista's average monthly profit (while dipping in the period immediately following GNN's exit) subsequently increased in the period from March to December 2010, and was higher than during the period of the complaint; and that (ii) there was a highly likelihood of recoupment by Forum under a range of assumptions. The Tribunal also noted that there had been no entrants into the market by the time of its judgment. In relation to whether GNN was an equally efficient competitor the Tribunal commented on various submissions in this regard, but ultimately concluded that GNN was not inefficient. According to the Tribunal's judgment, GNN was well run and well regarded in the local community newspapers market by both readers and advertisers, which appreciated its quality over its approximately 10 years of operation. It also won about one quarter of the market (making it larger than Forum). Media24 was suggesting that GNN was an inefficient competitor and as a result its exit would not cause consumer harm.

¹⁰² We note that this manager's testimony was considered as controversial. As the Tribunal wrote: '[this manager] was dismissed from his employment with Media 24 in April 2008 for allegedly making false travel claims. Criminal charges were levelled against him and he was convicted on some counts [...]' (para. 384) However, the Tribunal noted that it 'only accepted his testimony in respect of direct intent to the extent that it was consistent with documents from the record at the time, or where it has been corroborated by other witnesses' (para. 389).

1. Forum's profitability: Forum was consistently loss-making during the period of the alleged predatory conduct.
2. Forum's exit: following GNN's exit in April 2009, Forum's activities were downsized in mid-2009 and later in 2009 Media24 decided that Forum would be closed (the closure occurred in January 2010).¹⁰³
3. Forum's cannibalisation of Vista's advertising revenues: while, as set out above, the Tribunal could not reliably conclude on the magnitude of such effect, the Tribunal was comfortable that it was 'substantial' (para. 482).

In sum, based on the evidence summarised above, the Tribunal ruled that Media24 had engaged in exclusionary conduct through Forum's commercial practices and was in breach of Section 8c of the Act.

Comments We find this case instructive in a number of ways.

First, the taxonomy in the Act explicitly allows us to distinguish between: anti-competitive behaviour due to (predatory) pricing below average variable cost by a dominant firm; and anti-competitive behaviour due to more general exclusionary conduct, which the Tribunal accepted could occur as a result of a dominant firm pricing consistently below average total cost, with the intent of driving a competitor out of business.

This taxonomy broadly mirrors the *AKZO* rule discussed earlier in this chapter. The corollary to it is that an examination of evidence on intent is fundamental for the assessment of any predatory claims where prices are above average variable costs (though below average total costs). Media24 constitutes a case where the Tribunal appeared to have devoted significant effort in examining such evidence.

Second, the 'cannibalisation effect' set out above raises an important question on the treatment of opportunity costs (that is, in this case, foregone revenues) in the context of the calculation of avoidable costs following allegations of predatory conduct. As the Tribunal noted, one could criticise the inclusion of opportunity costs on the basis that any business decision that was not profit-maximising may lead to a dominant firm being found guilty of predation; and this may generate a chilling effect on competition. The Commission's economic expert responded to this critique by stating that he was not claiming that the opportunity

¹⁰³ As for why Forum was not closed immediately after GNN's exit, the Tribunal stated at paragraph 432: 'the delay in closing Forum was driven by two strategic considerations – the possible re-emergence of another competitor and legal concerns given the Commission's investigation.'

costs incurred by the dominant firm were about a failure to maximise profit by not charging some hypothetical price. Rather, he argued that the relevant benchmark was the ‘clear and obvious alternative that the dominant company would have pursued had it not pursued the allegedly abusive strategy’ (para. 150). This position seems sound to us as a matter of economics and policy, along with the need to consider opportunity costs carefully in a case such as this.

As for the theory of harm providing a rationale for predation, a potential candidate may be a reputation-based theory of harm. The Tribunal briefly considered this in its discussion of recoupment:

Academic writing and case law has also recognised what is termed ‘reputational recoupment’ or leveraged recoupment. Where the dominant firm operates in multi-geographic markets, as is the case with Media 24, then it benefits in its other markets from a reputational effect – i.e. being seen as a robust competitor. The reputational recoupment effect is thus a multi-market benefit from predation and not a benefit recouped in a single market. On the facts of this case Media 24, which has numerous community newspaper titles throughout the country, would have benefited in this way as well. (Para. 509)

The judgement also contains elements which are reminiscent of the ‘deep pocket’ theory of predation. For instance, at paragraphs 276–8 and 296–8, the Tribunal reports evidence according to which Media24’s managers seemed to be well aware that their strategy was inflicting losses on GNN (since Forum was not profitable, it was reasonable to assume that GNN, lacking Media 24’s economies of scale and scope, could not be making money). In fact, they seemed to be surprised that their rival paper lasted as long as it did.

More generally, throughout the judgement there appears to be the idea that Forum (a smaller and less profitable publication) was used as a ‘fighting brand’ to protect the market position of Vista. For instance, at paragraph 383 the Tribunal states: ‘The evidence of direct intention, whilst comprising some bellicose language, constitutes more than just war talk by a competitor about its rival. The strategy devised at Moolmanshoek and which is evidenced later in the documents, constitutes a plan to eliminate a competitor in the market by using one of its titles, which they had earlier on considered closing, as a barrier to expansion in the market to protect the market position of the larger title.’

1.6.9 ECS/AKZO

In what is probably the best known European case on predation, the European Commission found AKZO guilty of predatory pricing in the

market for organic peroxides, a chemical product used as a flour additive in the UK and more generally in the chemicals' industry. The Court of Justice upheld. According to the Commission, AKZO started to prey upon its smaller rival ECS when this (previously limiting itself to sell organic peroxides to be used as a flour additive in the UK) started to target a bigger market (organic peroxides for the polymer, or plastics, industry). ECS' attempted expansion included making competitive offers to BASF, one of AKZO's biggest continental clients in the polymer industry. The following excerpt from the Commission's decision suggests the motive behind AKZO's strategy:

The policy documents found at AKZO UK and AKZO Chemie show an explicit link between AKZO's policy in the plastics market and its action in the flour additives market. This link had in fact been crucial to the threats made to ECS in December 1979. AKZO was concerned in the long run to protect its market position in the plastics market and the most effective method of achieving this in the case of ECS was to react in the smaller flour additives market which was of only marginal importance to itself but accounted for most of ECS's turnover. (Para. 48 of the decision)

AKZO was found to be dominant in the market for organic peroxides sold as a flour additive, with a steady market share in excess of 50 per cent. To further substantiate the allegations, the Commission's decision reports instances of AKZO making below-cost offers to ECS' most important business clients, with serious effects on ECS:¹⁰⁴

The value of ECS' flour additives sales in the United Kingdom had by 1984 declined to 70 % of its 1980 sales [...]. In effect the 'independents' and Allied Mills lost to AKZO UK accounted for almost one-third of its flour additive business in the United Kingdom. The general decline in prices of flour additives also involved a reduction in the margins on the business which ECS retained. In order to remain in business (says ECS) it was obliged to increase its bank borrowings substantially thereby incurring additional bank charges and interest. The lack of available funds also caused ECS to reduce its budget for research and development and to delay modifications to its plant intended to deal with new organic peroxide business. (Para. 50)

The facts in *ECS/AKZO* are potentially consistent with several of the theories we presented in this chapter, though the limited public information prevents us from making a conclusive statement on the full applicability of any of them.

¹⁰⁴ The Commission also found some documentary evidence of a predation plan.

First, the above excerpt refers to facts that seem to be broadly consistent with the financial market predation discussed in Section 1.2.2.1 (by reducing ECS' profits, AKZO forced it to rely on more costly acquisitions of finance thereby limiting its ability to expand).

Second, there may have also been a reputation element in *ECS/AKZO*. According to the Commission (para. 48), the incumbent had already forced out of the market a competitor (Scado), which had tried to fight AKZO on its key customers and markets. The strategy vis-à-vis ECS looked similar - AKZO may have wished to build (or maintain) a reputation as a fighter against existing or new competitors that began attacking AKZO's key markets:

[T]he elimination of ECS would have a dissuasive effect upon any other small producer which might be minded to attack AKZO's established market position. (Para. 86)

Finally, at least in principle, this case may have featured the mechanism we explained in Section 1.3.2.6: with common costs (or more generally scope economies) to expand operations in the two markets (both involving organic peroxides), AKZO might have preyed upon ECS in its key market (flours' additives to UK and Irish customers) to deprive ECS of the scale it needed to expand also in the plastic market and thus allow AKZO to maintain its dominance across a broader range of markets. However, publicly available information prevents us from understanding how ECS (and not AKZO) understood the strategic link between the two markets, that is, the significance of potential scope economies.

1.6.10 American Airlines

Next, we consider a fairly recent US case, *American Airlines*. As mentioned in Section 1.5.1, where we briefly reviewed the history of the US case-law on predation, the District Court in *American Airlines* dismissed the case mostly on the basis that the Department of Justice ('DOJ') had failed to show that the defendant had priced below a reasonable measure of cost (and the Appeals Court upheld). Nevertheless, it is interesting to look at the economics of the case in more detail.

The complaint referred to American Airlines' ('AA') behaviour at one of its hubs (Dallas-Fort Worth, or 'DFW') over the period from 1995 to 1997. AA's strategy, which was a response to the pricing strategies of three low-cost carriers serving the hub, included fare reductions and capacity expansion over four routes (scheduling of extra flights, using larger planes

and/or making low fares available more frequently on these routes). At the end of the allegedly predatory period, the low-cost carriers abandoned certain routes where they were competing with AA or even exited the industry altogether. AA then realigned its prices to the pre-price war period and reduced its frequency in certain markets.

AA served around two-thirds of the passengers at DFW; its costs per available seat-mile, however, were almost twice as high as those of a competing low-cost carrier (mostly due to stringent trade union contracts), according to Edlin and Farrell (2004). This handicap was offset by economies of scope arising through the hub-and-spoke network at DFW which, as Edlin and Farrell (2004) – providing an extensive commentary of this case – noted, was the key mechanism for AA's success. The District Court too explained this quite clearly:

Operation of a hub, like American's at DFW, provides economies of traffic density that lowers the costs on a per-passenger basis and/or permits the hub operator to increase frequency. (Page 5 of the District Court's judgment)

Moreover, Edlin and Farrell (2004) referred to a DOJ Appellant's brief where there was a suggestion that AA began to worry that the low-cost carriers, if successful, would expand and set up a 'mini-hub' at DFW.

AA's incumbency advantage arose through its historic presence at DFW (serving over two thirds of passengers), with stable, high market shares on several routes. As the District Court noted:

On a number of nonstop routes from DFW, American had market shares ranging from 60% to 100% [...] for the period from 1990 to 1999. (Page 10)

The incumbency advantage materialised more broadly:

Frequency dominance or origin point presence advantages are reinforced by marketing programs including frequent flyer programs and travel agent commission overrides. American [Airlines'] investment in establishing its DFW hub involved a large sunk investment, and another airline with similar cost structure would also have to make large investments to build a similar hub at DFW. (Page 11)

The above elements are consistent with a plausible theory of harm based on the model we presented in Section 1.3.2.1. By preventing an entrant from serving a critical mass of (early) consumers, the incumbent may have successfully prevented the entrant's expansion, or even excluded the entrant altogether, as this could not build a network of a critical size, thus forgoing crucial economies of scope (and scale). The fact that AA worried about the entrants setting a mini-hub at DFW (its own hub) set this case apart from

Spirit Airlines, in the context of the model we proposed. In *Spirit Airlines* (which we discussed in Section 1.5.1), the entrant wanted to expand its own hub at Philadelphia, by providing two more routes (to Boston and Detroit-Metro); but this entry strategy was fiercely fought by Northwest, one of whose hubs was at Detroit-Metro, not Philadelphia.

Some commentators have argued that the case may be interpreted in the spirit of the reputation models we discussed in Section 1.2.2.2. Edlin and Farrell (2004) pointed out how the DOJ indeed put forward a theory of harm based on AA wishing to build a reputation as an airline that would fight entrants at its hub. AA itself – the District Court noted – conceded that there is a strategic element in the interaction between low-cost carriers ('LCCs') and the incumbent:

American [Airlines] believes that LCCs engage in 'game-theory' analyses when determining whether to enter, expand in, or remain in, a market in competition with an incumbent. American [Airlines] believes that if it permits an LCC to fly one flight in a market, that LCC will increase its frequencies and become a powerful competitor, and believes that it is valuable for competitors taking note of American [Airlines'] actions. (Page 73)

In addition, Kim (2009) used data from the period and routes that were broadly relevant for the *American Airlines* case, in the context of a reputation model where potential entrants observe the incumbent's actions, update their beliefs about the incumbent's 'type' (that is, soft or aggressive) and then decide whether to enter. He found some evidence supporting the (reputation) model he proposed.

However, the District Court strongly rejected a reputation-based predatory theory of harm (and the Appeals Court agreed).

The District Court also rejected the notion that AA could have made a recoupment, referring to the high bar set by *Brooke*. It stated that:

the uncontroverted evidence establishes that [Dallas airport] routes are not structurally susceptible to the supra-competitive prices which is a prerequisite to a successful predatory pricing scheme. That is, the nature of the relevant [Dallas airport] airline routes demonstrates that recoupment is not a dangerous likelihood. No hub airport in the country is served by more [low cost carriers] than [Dallas airport]. (Page 121)

There was some significant dispute between the parties on the exact price-cost test to apply, and which cost categories should be included in the correct cost benchmark for the price-cost test (see Edlin and Farrell, 2004, for a detailed account). In short, the DOJ performed four versions of price-cost tests, for each route under investigation, testing:

1. Whether AA's incremental cost exceeded incremental revenue: the DOJ compared the route-level profits after the capacity expansion to the route-level profits prior to the capacity expansion (but following the reduction in AA's fares that followed entry on each route) – ('Test One').
2. Whether AA's long-run average variable cost exceeded price: the DOJ adopted an 18-month period for the definition of long-run – ('Test Two').
3. Whether price was persistently (defined as for longer than 12 months) below AA's internal accounting cost measure used to measure profitability – ('Test Three').
4. Whether AA's incremental cost of a capacity expansion exceeded AA's prices – ('Test Four').

The DOJ submitted that the results of these tests showed evidence of profit sacrifice, consistent with a predatory strategy. AA retorted that, for each route under investigation, its prices were consistently above the average variable cost level.

The DOJ's Test One was dismissed by the Courts as they argued that it was a test of short-run profit-maximisation, which compared the route-level profits before and after the capacity expansion. The Appeals Court stated that '[s]uch a pricing standard could lead to a strangling of competition, as it would condemn nearly all output expansions, and harm to consumers' (para. 23 of the Appeals Court's judgment).

Both the District Court and the Appeals Court dismissed the DOJ's Tests Two and Three on the basis that the measure of AA's internal accounting cost used for these included some categories of fixed costs. And since the Appeals Court held that the correct cost benchmark to use was marginal cost (and that average variable cost was generally a good proxy for it, albeit not necessarily the only one), the measure of AA's internal accounting cost used for Tests Two and Three was therefore an inappropriate proxy for marginal (or incremental) cost.

Finally, Test Four was dismissed on the basis that both Courts found that the measure of AA's internal accounting cost used for this test did not solely capture AA's avoidable costs of the incremental capacity expansion on a given route, but also some general common costs (that is, costs that were not specific to that capacity expansion).

Overall, it is not clear based on the information available in the public domain that in this case there was sufficiently robust evidence of prices below the relevant cost benchmark to support a predatory finding.

In sum, *American Airlines* is an interesting case in the sense that it broadly matches the features of an economically sound theory of harm (based on scope economies), and perhaps of a reputation-based one, but where the evidence on whether a price-cost test is met is mixed.

1.6.11 Standard Oil and American Tobacco

We conclude this section with a discussion of two very early cases in US antitrust, *Standard Oil* and *American Tobacco*, which we already touched upon in Section 1.5.1. We provide a joint commentary, as their facts were strikingly similar and so were the economic mechanisms at work.¹⁰⁵ Further, differently from the previous discussion of other cases, we draw more extensively from other authors for our discussion of these cases. This is due to the fact that these landmark US cases are quite dated and have attracted significant research and commentary, which we believe to be insightful and worth reporting.

The Standard Oil company was born in Ohio in 1870 through the merger of three refining businesses. The newly formed entity then started expanding through subsequent acquisitions of refineries in other States, until it reached (by 1882) a national market share of 90 per cent, in the businesses of refining, transporting and retailing petrol. Its business model was based on obtaining preferential rates from railroad companies, closing many of the plants bought and putting certain plants under control that was only apparently separate from the parent company. After the turn of the century, Standard Oil further resorted to other practices, including selective price cuts and organising a market sharing scheme among its 37 affiliates across the United States. The Supreme Court condemned these practices as both restraints of trade and unlawful monopolisation.

The American Tobacco Company was also formed after a merger between (five) competitors. It was active in the cigarettes and tobacco markets. The company then grew through subsequent acquisitions, often closing plants, reaching a national market share of 86 per cent in domestic cigarettes by the turn of the twentieth century. The few competitors who did not oblige were fought off with very strong price cuts, some leading to prices below cost (in the case of plug tobacco American Tobacco lost more \$4m, at that time's prices). By 1904, American Tobacco had purchased (and closed down in certain cases) over 30 companies.

¹⁰⁵ In what follows, we also draw from Evans (1912).

As we already noted in Section 1.2.1, McGee (1958) remained unconvinced of the use of predatory pricing in *Standard Oil*. His main argument was that engaging in a price war would have simply been irrational, as the alternative of buying the competitor off and starting earning monopoly rents immediately was more profitable, an argument espoused more generally by Telser (1966).¹⁰⁶ McGee complemented his logical arguments with an analysis of the records of the proceedings (including testimonies) and concluded against the finding of predatory pricing, without, though, any econometric evidence as such.

McGee admitted that, in principle, an intimidating strategy revolved around predatory pricing could yield lower purchase prices and thus be overall beneficial to the incumbent.¹⁰⁷ Yet, he argued that the predation-through-merger hypothesis failed in the case of *Standard Oil*.

Dalton and Esposito (2007) re-examined the judgment records of four of the five ‘major’ cases considered in McGee (1958) and drew the conclusion that *Standard Oil* had actually engaged in predatory pricing. The authors distinguished between predation and competition on the merits (referred to as ‘aggressive pricing’) on the basis that both entailed low prices, but the latter caused supply to expand and prices to remain low in the long run, whereas the former had the opposite effects. The authors complemented their analysis with a price comparison against *Standard Oil*’s own costs and with evidence on the intent to predate where these elements were available in the records; both tests supported their conclusion.

Burns (1986) set out to test the predation-through-merger hypothesis in the context of *American Tobacco*. He went even further, postulating that predatory pricing against a competitor may generate further benefits – in terms of lower purchase prices of subsequent preys – through a *reputation effect*. Using a sample of 43 competitors bought out by *American Tobacco* at the turn of the twentieth century, Burns found evidence consistent with his rational predation hypothesis: predatory pricing had a significant downward effect on the purchase price of the prey in question (up to 60

¹⁰⁶ Yamey (1972) – writing more generally, and not specifically about this case – also suggested that a more profitable alternative to predatory pricing might be the formation of a cartel with some entrants.

¹⁰⁷ Recall from Section 1.2.2.3 that the argument behind a predation-through-merger strategy runs as follows: by setting ‘low’ prices, an incumbent could (profitably) give a (false) signal to an entrant that the market conditions are structurally bad; the entrant would therefore expect its future cashflows to be low and in turn accept that its valuation as an enterprise should likewise be relatively low (since the value of a firm is the present discounted value of its expected future profits). With this conjecture, the entrant should be willing to be taken over for a low price, to the benefit of the incumbent.

per cent of what they would have cost absent predatory pricing) as well as an additional indirect downward effect (of around 25 per cent) on the acquisition price of other preys (which can be interpreted as a reputation effect). These findings would be consistent with the interpretation of *American Tobacco* also within the framework of the reputation model described in Section 1.2.2.2 (with the minor difference that the original models entail an entry versus no entry decision, whereas in *American Tobacco* the corresponding choice variable was fighting versus selling out to the incumbent).

In sum, both *Standard Oil* and *American Tobacco* presented features consistent with both a theory of harm based on signalling (via predation-through-merger) and with one based on reputation effects.

Price Discrimination and Single-Product Rebates

2.1 Introduction

Chapter 1 deals with predatory pricing, that is, typically, low prices offered by a dominant firm *across the board*, to all of the customers who are buying a certain product at a given point in time. In this chapter we discuss, instead, the possible exclusionary effects of price discrimination in its various forms (including different types of rebates and discounts). We also suggest, based on the economic theories reviewed, a possible approach that competition authorities may want to follow when considering instances of potentially anti-competitive rebate schemes. Differently from predation, in this chapter we focus on low prices offered to specific buyers, or for specific units demanded by buyers. Some forms of price discrimination may also be conditional on buying different products, but we shall deal with bundled discounts in Chapter 4.

This chapter proceeds as follows. In Section 2.2 we first define price discrimination and discuss its welfare effects in general, that is, when exclusion is not an issue. In Section 2.3 we study the circumstances under which a dominant incumbent firm may use price discrimination (in its different forms) to exclude a rival, and show that the more individualised and targeted the discrimination the more likely that it will have exclusionary effects all else equal. In technical Section 2.4 we formalise the analysis of exclusionary discrimination. In Section 2.5 we draw policy implications from the theory. In Section 2.6 we discuss key decisions by competition authorities and landmark case-law. Finally, in Section 2.7, we discuss a few antitrust cases investigated in different jurisdictions and seek to interpret them in light of some of the models reviewed in this chapter.

2.2 Price Discrimination, Welfare and Efficiencies

2.2.1 Forms of Price Discrimination

Price discrimination consists of different consumers paying different unit prices for the same good, when it costs the firm the same amount to produce and serve these consumers (or, more generally, when prices are at different ratios to marginal costs).¹ It is a very widespread phenomenon and in practice it may take different forms, as we shall discuss below.

Economists typically distinguish three types of price discrimination. *First-degree* price discrimination refers to a theoretical situation in which a firm knows exactly each consumer's valuation (or willingness to pay) for its product and charges her the price which equals her valuation, thereby extracting all her surplus. Under *second-degree* discrimination, the firm cannot explicitly discriminate across consumers (either because it cannot observe them or because it is not allowed to offer different deals to *different* consumers). It therefore offers the same menu of prices to all of them, trying to induce them to 'self-select' and choose a particular deal. For instance, an airline may not be able to observe who flies for leisure and who flies for business (the former would typically have a higher price elasticity of demand than the latter, that is, they would be more price-sensitive), but offers a large discount for a return date which involves spending the weekend at the destination – thereby inducing business-people (who generally want to return home for the weekend) to pay a higher price for their ticket. Quantity discounts may be another instance of second-degree price discrimination: all consumers are offered the same schedule whereby the unit price falls with the number of units bought, but large buyers will end up paying a lower unit price than small buyers.² *Third-degree* discrimination occurs instead when a firm charges different prices to groups of consumers having different (observable) characteristics. For instance, a firm may expect that people below 21 years or above 65 years may be still studying or respectively be already retired, and therefore have a higher price-elasticity, and may want to offer them discounts.

In all these cases, the objective of the firm is to extract the highest possible surplus from consumers, by making them pay as much as they

¹ For a discussion of the definition of price discrimination, see among others Varian (1989: 598) and Tirole (1988: 134). For a thorough review of models of price discrimination, see Armstrong (2008) and Stole (2007).

² Two-part tariffs, composed by a fixed fee and a linear price component, can be interpreted as a form of quantity discounts because the unit price decreases with the number of units bought.

would be willing to. Such a strategy crucially depends on consumers not being able to engage in *arbitrage* (that is, trade between them in a mutually beneficial way). Otherwise – referring to the examples above – a low valuation consumer, a large buyer, or a person not in a working age would buy at a low price and resell it to other consumers (with higher valuation, with lower elasticity, or in working age) at a higher price.³ For these reasons, firms often try to prevent arbitrage, by imposing contractual limitations to buyers (for instance forbidding resale of their products) or conditioning purchase to some conditions (for instance by requiring a consumer to show proof of identity to enjoy a given service at lower price).

Price discrimination (across customers, or across the different units that a given customer purchases) can be implemented in different ways. For instance, a firm may offer *selective price cuts*, that is, discounts to specific customers that the firm is able to identify. Alternatively, a firm may offer ‘coupons’, that is, tickets that can be redeemed for a discount when purchasing a product. This may allow them to target the discount to customers whose demand is more elastic because these are the ones that actually use the coupon. Another way to implement price discrimination, which is common especially in intermediate-goods markets, consists of offering a discount conditional on fulfilling some conditions.

For instance, *quantity discounts* are rebate schemes that grant a discount under the condition that the customer’s purchases exceed a given quantity threshold within a given period of time. If the threshold to qualify for the discount is expressed in terms of percentage increase over the previous period’s sales, the rebate scheme is denoted as *dynamic* (or growth target) rebate.

Market-share discounts represent another recurrent type of conditional rebates. In this case, the discount is based on the supplier’s share of the overall customer’s purchases. A market-share discount which is conditional on buying most (or all) of the input requirement from a specific supplier can be denoted as a *loyalty* discount (or exclusivity discount).⁴

If the buyer receives a rebate only on the units exceeding the target threshold, the discount is denoted as *incremental* rebate. If instead, once

³ Arbitrage across countries is often called *parallel trade*.

⁴ The definition may change across agencies, judges and scholars. For instance, according to Elhauge (2009), loyalty discounts are agreements to sell at a lower price to buyers who buy *all or most* of their purchases from the seller. Instead, the General Court in *Intel* denotes such discounts as *exclusivity discounts*, whereas loyalty discounts (or fidelity discounts) are those discounts that are not conditioned to exclusive or quasi-exclusive supply but in which the mechanism for granting the discount has a fidelity-building effect (see Sections 2.6 and 2.7 for a more detailed discussion of *Intel*).

the target threshold is reached, the buyer receives the rebate on all the units previously purchased within an agreed period, the discount is referred to as *retroactive* rebate (or all-units discount).

Furthermore, rebates are denoted as *standardised* if the same scheme is offered to all the buyers, for instance a quantity discount with the same target threshold to qualify for the discount. Rebates, instead, are denoted as *individualised* if the scheme differs across buyers.

To conclude, a note of warning. Although it may appear that a discount conditional on buying 100 per cent (or most) of the buyer's requirement is equivalent to an exclusive dealing contract, the two differ in an important dimension. Exclusive dealing contracts are long-term *bilateral* contracts that involve a *commitment by the buyer* not to purchase from alternative suppliers during a given reference period. This commitment component on the side of the buyer is not present in loyalty discounts, which are *unilateral* offers in which it is only the supplier that commits to offer different terms of trade depending on how much the buyer purchases. Differently stated, a buyer that enters an exclusive dealing contract with a supplier cannot purchase from another supplier. Instead, in the case of exclusivity discounts, a buyer can switch at any moment to an alternative supplier, even though it will obviously lose the discount. As we discuss in Chapter 3 (which focuses on exclusive dealing), the commitment on the side of the buyer may be a crucial factor for the incumbent to use long-term contracts and exclude a more efficient rival. (See Ide et al., 2016.)

Having defined price discrimination, we next discuss its welfare effects, abstracting from its possible exclusionary role. This will allow us to highlight that the welfare effects of price discrimination are ambiguous in general, something that will turn out to be crucial when discussing policy implications.

2.2.2 Welfare Effects of Price Discrimination

2.2.2.1 Price discrimination towards final consumers

For some reasons – perhaps because firms resort to it in order to make customers pay a price closer to their valuation, perhaps because the very word ‘discrimination’ may sound in conflict with principles of equality – many people feel that price discrimination is a harmful practice. In fact, economic theory shows that in general (recall, here we discuss price discrimination by *any* firm and abstract from exclusionary motives, which we shall address later in the chapter) the welfare consequences of price discrimination are ambiguous.

Let us see an example of why, in general, welfare effects are ambiguous.⁵ Suppose a monopolistic firm sells the same product to two different groups of consumers, call them rich (R) and poor (P), with the demand of rich consumers being less reactive to price changes than the demand of poor consumers.⁶ If price discrimination is allowed, then the firm will find it optimal to set a higher price for consumers whose demand is less elastic: $p^R > p^P$. If instead price discrimination was banned, the monopolist would have to set a uniform price, p^U . How will the firm choose this uniform price? There are two cases to be considered.

(1) The firm may find it optimal to serve both groups of consumers. In this case, it will set the price p^U somewhere between p^P and p^R . By comparing this situation with that of uniform prices, one can conclude that when price discrimination is allowed profits will increase, because the firm is able to charge consumers according to their price elasticity, rich consumers will enjoy a lower surplus (they buy at a price p^R which is above p^U), whereas poor consumers will have higher surplus (the price p^P is below p^U). These effects have opposite signs and their sum is *a priori* ambiguous. The general rule is that price discrimination decreases welfare if it does not increase total output.⁷

(2) The firm, if forced to choose a uniform price, may find it optimal to set $p^U = p^R$, thereby serving rich consumers at their optimal price and not supplying poor consumers (or not supplying most poor consumers if some of them still buy at the price $p^U = p^R$). This might be the case if the proportion of poor consumers in the market is small and the profits lost by not serving them are outweighed by the high profits on rich consumers. Comparing this situation with that of price discrimination, one would now conclude that allowing price discrimination would leave rich consumers with the same surplus, while it would increase the surplus of the poor (as they would not buy, or most of them would not buy under price uniformity) and the profits of the firm. In this case, banning price discrimination would decrease total welfare.

Dynamic effects of price discrimination: incentives to invest Price discrimination might also affect dynamic efficiency, through the firms'

⁵ For simplicity we consider here the case of a monopolistic supplier, but the same conclusion holds in the presence of oligopolistic sellers. See Corts (1998) and Armstrong (2008).

⁶ This is a case of the so-called third-degree price discrimination.

⁷ See Schmalensee (1981) for a seminal contribution on this issue. See also Varian (1989), Tirole (1988:137–8), Motta (2004: sections 7.4.1.3 and 7.4.2), and more recently, Aguirre et al. (2010). Cowan (2012) shows that, even if discrimination reduces total welfare, consumer surplus can rise under reasonable conditions.

incentives to invest. This is because a firm's investment decisions will depend on the additional profits that the firm expects to make by investing. To the extent that price discrimination increases the marginal profits of the investment, it will also increase the amount of investment the firm will want to make.⁸ Through this channel, therefore, price discrimination might be welfare-beneficial. It would be very difficult, though, to verify and quantify these efficiency gains in concrete cases.

2.2.2.2 Price discrimination in input markets

So far we have considered price discrimination towards final consumers. It has been argued that price discrimination has additional negative effects when an upstream supplier sells an input to downstream firms. Since the (derived) input demand of a downstream firm is less elastic if its own final good is more attractive to consumers or if it is more efficient, by analogy to the case of price discrimination in final good markets, the supplier should then optimally charge the more efficient firms, or firms with more appealing products, a higher wholesale price. (See De Graba, 1990 and Yoshida, 2000.) A ban on price discrimination, by leading to an intermediate wholesale price, would then decrease the wholesale price for the more efficient firm and increase it for the less efficient firm, thereby leading to more production being undertaken by the more efficient firm and increasing total welfare. Further, a ban on price discrimination would remove the disincentive to invest in cost-reducing technologies that exists under discrimination, due to the fact that the benefits of becoming more efficient are mitigated by the prospect of paying a higher wholesale price.

A recent literature has shown that the previous insight may be misleading. For instance, while De Graba (1990) and Yoshida (2000) restrict the upstream firms to use linear prices when trading with downstream firms, Inderst and Shaffer (2009) allow for two-part tariffs.⁹ Under two-part tariffs the upstream supplier can use the fixed fee to extract the surplus of downstream firms, and the wholesale prices to maximise the overall industry profits. To do that, the upstream supplier will optimally choose a lower wholesale price for the more efficient downstream firm, thereby amplifying, rather than dampening, their efficiency differences. Consider

⁸ For a formalisation see Motta (2004: section 7.4.2.3). Inderst and Valletti (2009) study the effect of price discrimination on investment incentives when discrimination concerns input prices. We will discuss this paper in the next section.

⁹ All of these papers assume that the supplier can make take-it-or-leave-it offers and that the demand functions for the final products are linear.

now the case where price discrimination is banned so that the fixed fee and the wholesale price need to be uniform across buyers.¹⁰ First, the fact that the wholesale price must be uniform will lead to a price increase for the more efficient downstream firm. Second, now the (uniform) fixed fee will allow the supplier to extract the entire surplus of the less efficient downstream firm whereas some rents will be left to the more efficient firm (namely the additional profits it makes, given the same wholesale price, as compared to the less efficient firm). This will induce the supplier to increase the wholesale price further in order to minimise the rents left to the more efficient firm.¹¹ Inderst and Shaffer (2009) show that, under linear demand, a ban on price discrimination leads to higher wholesale prices for both firms, thereby reducing total welfare on two accounts: it increases the dead-weight loss to society due to higher final-good prices and it shifts a larger share of the now smaller total output to the less efficient downstream firm.

A ban on price discrimination will be welfare-detrimental also when downstream firms cannot observe their rivals' contracts. In this environment, the upstream supplier has an incentive to engage into opportunistic behaviour with its downstream firms, which can lead to extreme intrabrand competition and low profits (Hart and Tirole, 1990; O'Brien and Shaffer, 1992; McAfee and Schwartz, 1994. See the discussion in Section 2.3.3.1). The intuition is the following. Imagine that the upstream supplier offers downstream firms a pair of contracts that induce the joint-profit maximising outcome. For instance, the upstream supplier can offer contracts with a high wholesale price which induces downstream firms to set the monopoly price in the retail market. If one firm accepts it, the upstream supplier will have an incentive to offer a discount to the rival downstream firm, which induces it to produce more and increases their bilateral joint profit at the expense of the firm that has accepted the initial contract. Anticipating this, no downstream firm will accept the initial contract offers. It can be shown that the contracts that are accepted in equilibrium by the downstream firms involve higher quantities than the ones that maximise joint industry profits and lower final prices. A ban on price discrimination, instead, effectively removes any temptation from the monopolist to offer a price discount to a buyer after having signed with another (since a discount should be offered to all buyers), and therefore provides it with a commitment mechanism

¹⁰ However, (uniform) two-part tariffs still entail some degree of price discrimination as the unit price varies with the number of units that are bought.

¹¹ The intuition is that the more efficient firm will be affected on a larger volume base when the common wholesale price increases and thus the additional profits it makes will decrease.

that allows it to restore all of its market power and to sustain the joint-profit maximizing outcome. Hence, a ban on price discrimination leads to higher final prices and lower total welfare.

Finally, a ban on price discrimination might reduce total welfare and consumer surplus also by stifling incentives to invest and innovate, as shown by Inderst and Valletti (2009). The paper assumes that the upstream monopolist can only use linear prices. Downstream firms, by paying a fixed cost, can have access to an alternative (but inferior) supply option. It turns out that the upstream supplier will offer a lower wholesale price to a more efficient downstream firm since this finds the alternative supply option more attractive. When downstream firms make their investment decisions in a cost-reducing technology, each of them anticipates that by becoming more efficient than rivals it will also benefit from the subsequent larger input price discount. Then, downstream firms will have stronger incentives to invest relative to the case where uniform prices are imposed.

To conclude, these papers seem to suggest that, where discrimination involves input prices, the concern about possible welfare-detrimental effects is even less justified than in the case where discrimination involves final prices.¹²

2.2.2.3 Single-product conditional rebates: efficiency justifications

In this section, we discuss a number of efficiency justifications for these rebate schemes, which may explain why their use is so widespread by both dominant and non-dominant firms.

A common justification for conditional rebate schemes (applied on individual transactions) is that, under the presence of lumpiness in logistic or transportation costs, they allow a manufacturer to minimise such costs by inducing retailers to make sufficiently large orders. Further, conditional rebates (applied on the entire demand of a buyer during a reference period) can allow manufacturers to provide better incentives to retailers in relation to their selling efforts, which may be challenging to define effectively in a contract. For instance, an exclusivity discount, by inducing the retailer to purchase most of its requirement from the supplier that offers the discount, may stimulate the retailer to focus its promotional activity on the suppliers' products, or it may motivate the retailer to provide brand-specific information or customer services. (See Mills, 2010, for

¹² In Chapter 5, we discuss how these results change in a context in which a vertically integrated firm sets input prices in a discriminatory fashion between its own downstream affiliate and competing downstream firms which are not vertically integrated.

an analysis of this role of market-share discounts under the presence of heterogeneous retailers.) The logic behind the efficiency justification in this case is very similar to the one concerning exclusive dealing contracts that we discuss in more detail in Chapter 3.

Another welfare-beneficial effect of quantity discounts or market-share discounts relates to the mitigation of the so-called double marginalisation. When both a supplier and its retailer have some degree of market power, the final price will reflect the margin that each of them applies over their own cost. This may make the final price higher (and the output lower) than the price that maximises joint industry profits. When the supplier uses a linear wholesale price, double marginalisation would be alleviated by decreasing the wholesale price, but this would also decrease the supplier's profits. The problem can be solved or mitigated by using non-linear pricing, such as quantity discounts, which provides incentives for the retailer to sell the optimal quantity (and set the optimal final price), and allows the supplier to obtain sufficiently high profits. Quantity discounts will then increase not only industry profits but also consumer surplus, as they cause retail prices to fall as compared to linear pricing.¹³

Conditional rebates can be a way to charge very low prices, possibly below costs, on marginal units. As discussed when analysing predatory pricing (see Chapter 1), such a pricing can be rational for a firm, absent any exclusionary strategy, in particular situations in which additional sales provide side-benefits, for instance by increasing consumer goodwill, by allowing learning-by-doing, by producing network effects or by increasing the demand for complementary products (for example, in two-sided markets).

Further, as with price discrimination in general, conditional rebates may promote firms' investments to the extent that they increase the marginal benefits from investment, and will increase total welfare also through this channel. The same consideration made above, about the difficulty in assessing and quantifying these gains in concrete cases, applies also here.

A number of recent contributions focus, instead, on the role of conditional rebates as a screening device for a manufacturer when retailers have private information about consumers' demand. (See Kolay et al., 2004 and Majumdar and Shaffer, 2009.)¹⁴ A common theme is that rebate

¹³ Kolay et al. (2004) demonstrate formally how different forms of quantity discounts, such as two-part tariffs, incremental rebates and all-unit rebates, can equivalently solve the double-marginalisation problem under complete information about consumer demand.

¹⁴ These papers consider the case of a monopolistic supplier. Kolay et al. (2004) compare the use of all-unit discounts to that of incremental discounts. Majumdar and Shaffer (2009) compare market-share discounts to quantity discounts.

schemes may allow a manufacturer to introduce fewer distortions in order to extract informational rents from retailers, and for this reason under some circumstances they may increase total welfare.

By contrast, Inderst and Shaffer (2010) identify a potential anti-competitive effect of market-share discounts, not related to exclusion. The paper shows that market-share discounts allow a dominant supplier to increase industry profits by dampening *both* intrabrand and interbrand competition. Suppose that the retailers sell not only the dominant firm's product but also a substitute product. Relative to the price that a fully (horizontally and vertically) integrated firm would set, competition among the retailers would lead to lower prices. The dominant firm may want to relax retailer competition and raise prices by setting a higher wholesale price on its good. However, this relaxes intrabrand, but not interbrand competition: the retailers will have an incentive to sell more of the substitute product. By using market-share contracts, instead, and providing discounts which are conditional on the retailers selling a given proportion of its product, the dominant firm can both dampen competition on its own good and prevent the diversion of its sales.¹⁵

To sum up, from the above discussion one can conclude that, abstracting from exclusionary effects, there is no theoretical ground for systematic concerns against price discrimination. Rather, the main conclusion from economic models is that the welfare implications of price discrimination (or equivalently of banning it) are ambiguous.

2.3 Price Discrimination as an Exclusionary Device

We have dealt so far with the effects of price discrimination abstracting from its possible exclusionary effects. We next turn to the core issue and investigate whether price discrimination in its various forms may allow a dominant firm to exclude smaller rivals and harm consumers, an issue which has attracted relatively little attention in the theoretical literature. We first briefly discuss, in Section 2.3.1, the paper by Armstrong and Vickers (1993) where discrimination may deter entry for non-strategic reasons. Second, in Section 2.3.2, we move to the mechanism which is central to this book, by mentioning the important work by Innes and Sexton

¹⁵ If the dominant supplier tried to avoid diversion of its sales by offering quantity discounts, the final price of its product would end up being too low. The market-share discount is instead able to soften both interbrand and intrabrand competition, while alternative instruments such as wholesale price and quantity discount could only address one of these objectives at a time, while creating a conflict with the other.

(1993) who first emphasised the exclusionary role of ‘divide-and-conquer’ strategies, and then discussing more recent works (Karlinger and Motta, 2012 and some extensions of Fumagalli and Motta, 2013) which reflect the mechanism based on depriving the rival of the scale it needs. Third, in Section 2.3.3, we review some economic models that focus on contracts that reference (or benchmark) rivals;¹⁶ these are contracts whose terms depend directly or indirectly not only on the relationship between a seller and a buyer but also on the relationship between that buyer and the seller’s competitors (for instance, a market-share discount offered to a buyer would also depend on how much of its needs the buyer purchases from other sellers). Finally, in Section 2.3.4, we shall make a few remarks on a literature modelling the telecommunications markets which suggests that discrimination between ‘on-net’ and ‘off-net’ prices may exclude smaller telecommunication networks (Lopez and Rey, 2016).

2.3.1 Non-strategic Selective Price Cuts

In this section, we discuss the possible exclusionary effects of a firm which discriminates across the markets in which it operates. To motivate the issue at hand, consider the following example. Suppose that in a given country there exists a dominant incumbent firm that produces a good with sizeable transportation costs. This firm is located in the centre of a country whose population is concentrated around two provinces, one in the North and the other in the South. Suppose now that a rival sets up a plant in a neighbouring country which lies in the North. Intuitively, there is little doubt that the dominant firm will want to engage in price discrimination, and set a higher price in the Southern province, where it faces no competition, than in the Northern one. But the question is whether it should be allowed or not to price discriminate.

This simple example is inspired by an actual competition case, *Irish Sugar*, in which price discrimination was found to be unlawful by the European Commission (and then by the General Court), despite the post-rebate prices being above costs.¹⁷ In such a case, it could not be

¹⁶ To our knowledge, this terminology was first used by Fiona Scott Morton, in ‘Contracts that Reference Rivals’, Speech at Georgetown University Law Center, 5 April 2012.

¹⁷ The other key European case where above-cost selective price cuts were found abusive was *Compagnie Maritime Belge*. We shall return both to this case and to *Irish Sugar* in Section 2.6.2. For more policy and legal detail on above-cost selective price cuts, the reader is directed to Elhauge (2003) and Whish and Bailey (2015).

said that the firm was engaging in predatory pricing, so should price discrimination by an incumbent firm be treated as an independent abuse?

A paper by Armstrong and Vickers (1993) helps shed light on this question.¹⁸ Let us summarise the model as follows. The incumbent firm, call it firm *I*, is selling to two buyers, call them *S* and *N*, for ‘South’ and ‘North’, respectively. Buyer *S* is captive to the incumbent, while buyer *N* can be ‘contested’ by the rival firm *E*, if this firm decides to enter the market. (Equivalently, think of them as a market *S* in which the incumbent is monopolist and a market *N* in which it is a duopolist.) Suppose first that firm *E* is already in the market.¹⁹ If price discrimination was allowed, the outcome of the competitive process would be that firm *I* sets a high price to buyer *S* and a low price (but above its marginal costs) to buyer *N*. If price discrimination was banned, and firm *I* was thus obliged to set the same price to each buyer, it would set a (uniform) price which is between the discrimination prices: it is intuitive that under price uniformity, it is not optimal for firm *I* to choose a common price as low as the discriminatory price for buyer *N* as it would forgo large profits on the other buyer *S*. Hence, when discrimination is possible, firm *I* would make higher profits than in the case where it is restricted to set uniform prices (because it is free to raise prices to buyer *S* as much as it wants, while competing fiercely for buyer *N*), while firm *E* would make lower profits (since firm *I* is more aggressive on buyer *N*).

Suppose now that firm *E* is not in the market yet, but it is simply considering whether to enter (and pay a fixed entry cost *f*) or not. Firm *E* also knows if price discrimination is allowed or not and it is able to anticipate the outcome of the competitive game in case it entered. As a result, it may well be that with price discrimination it expects that its profits would be insufficient to cover its entry costs (because firm *I* would behave more aggressively), whereas if price discrimination was banned it would make enough profits to cover them. Therefore, price discrimination would lead it to stay out of the market.²⁰

¹⁸ See also Armstrong (2006b: section 4.3) for a stylised version of this model and for a more general review of models of price discrimination; and Motta (2004: section 7.4.2.5) for a simple version of a similar model to the one we discuss in the text.

¹⁹ Technically, we should also assume that firms compete in prices and that buyer *N* regards products as differentiated. The actual model of product differentiation chosen does not change the qualitative results.

²⁰ To be more precise: if *f* was sufficiently low (that is, lower than the profits made under price discrimination), then firm *E* would always enter. If *f* was sufficiently large (that is, larger than the profits firm *E* would make when discrimination is banned) it would never

In this sense, price discrimination would deter entry, and one may think it would be 'bad'. Before jumping to conclusions, however, a few remarks are due.

First of all, note that the dominant firm does not adopt any strategic conduct here. When it price discriminates, it does not do so in order to exclude the rival. In fact, we have seen that discrimination takes place when the rival is already in the market and there is no hope to make it exit. Price discrimination is simply the result of more intense competition in the market where a rival is active.²¹ As a result the incumbent does not set a price below costs and does not make any losses. It is something very different than in situations of predatory pricing, where the dominant incumbent engages in profit sacrifice in order to exclude rivals! It is accordingly difficult to say that firm *I* is behaving anti-competitively. (And we know from the previous section that there is no ground for banning price discrimination abstracting from exclusionary effects.)

Second, the welfare effects of banning price discrimination would be ambiguous. True, a ban on price discrimination may in some circumstances promote entry. However, there are at least two considerations that suggest that a ban might result in adverse welfare effects in some circumstances. (i) The ban may promote inefficient entry. Imagine that the entrant is less efficient than the incumbent (it has higher marginal costs). When there is a ban on discrimination, the incumbent will be less aggressive on the contestable buyer, and it will end up setting a (uniform) price well above its own marginal costs (because it does not want to forego too much profit on the captive buyer). As a result, even an inefficient rival would be able to survive in the market, giving rise to productive inefficiencies which decrease welfare (a more efficient seller is partially replaced by a less efficient one). (ii) Suppose that the government does not know the exact values of the fixed costs (or equivalently of the profits

enter. Discrimination would therefore deter entry only for the remaining intermediate values of f .

²¹ It is well known that price discrimination leads to more aggressive competition (since firms would compete on each customer group). See also Thisse and Vives (1988), who study a game where firms choose the price regime before competing: discriminatory prices would emerge as an equilibrium for a prisoner dilemma mechanism, but firms would be better off if they could jointly commit to uniform pricing. Corts (1998) shows that, in oligopolistic markets, price and welfare effects of price discrimination may differ according to whether firms have the same ranking of consumer groups in terms of demand elasticity or not. For instance, if like in the Thisse and Vives (1988) model, the strong market for a firm is the weak market for the other, then price discrimination is welfare-beneficial to consumers. If, instead, firms rank markets in the same way, price discrimination may harm all consumers.

under the different regimes) when it sets the policy. It may be that the ban allows the entrant to enter when price discrimination would have kept it out of the market. But it may also be that the entrant would have entered anyhow (because fixed costs f are below the profits it makes under price discrimination). In this case, a ban would have ambiguous consequences, as it would decrease surplus for the incumbent and for the contestable buyer (it would pay more under the ban) and increase surplus for the entrant and the captive buyer (it would pay less under the ban).

2.3.2 Price Discrimination to Deprive the Rival of the Scale it Needs

The model by Innes and Sexton (1993), which we will discuss in Section 2.3.2.1, is probably the first one to illustrate the potential exclusionary role of price discrimination, and to highlight a mechanism which would be central to the subsequent literature on exclusionary conduct: when entry entails scale economies, the incumbent can use a ‘divide-and-conquer’ strategy whereby a certain number of buyers are offered good terms of purchase, thereby reducing the scale available for entry and exerting a negative externality on the remaining buyers. This allows the incumbent to charge the remaining buyers a much higher price. The notion that an incumbent firm can exclude more efficient rivals by exploiting this externality is the key idea of this book. A similar externality is central to the model of predation (due to Fumagalli and Motta, 2013) that we discussed in Chapter 1. This externality is also central to the paper by Karlinger and Motta (2012) where selective price cuts can exclude a more efficient rival, reviewed in Section 2.3.2.2, and to the literature on anti-competitive exclusive contracts, which we discuss in Chapter 3. This externality also arises in some extensions of the main model discussed in this Section (which rationalises conditional rebates), and we discuss some of these extensions in Section 2.3.2.3 and 2.3.2.4. Finally, in Section 2.3.2.5, we discuss retroactive rebates and the price-cost test that can be adopted for these rebate schemes.

2.3.2.1 Divide-and-conquer price discrimination

Innes and Sexton (1993) consider an incumbent firm which makes price offers to customers. (See Section 2.4.1 for a formal analysis.) Customers who reject the incumbent’s offer can organise themselves and enter the market as producers of the good. Think, for instance, of consumers that could set up and operate a cooperative firm; or of intermediate goods markets, where buyers are often few and large, and can be credible entrants;

or of large retail chains which can start selling their private labels. Customer entry entails fixed set-up costs and organisation costs. These costs are high enough to make it unprofitable for a buyer alone to enter upstream production, even if the incumbent offers the monopoly price. Instead, two buyers together sharing these costs would find it profitable to engage in self-supply, if the incumbent offers the monopoly price.

In this environment, imagine that the incumbent offers the monopoly price p^m to one buyer and a much lower price p^* to the other. The price p^* must be low enough to make a buyer prefer the incumbent's offer to self-supply with the other buyer.²² Hence, a buyer that is offered p^* accepts for sure the incumbent's offer, irrespective of the choice of the other buyer. Anticipating this, the buyer that is offered the monopoly price has no better option than accepting as well, since entering upstream production individually is not profitable. The above 'divide-and-conquer' strategy turns out to be successful to discourage customer entry.

Consider now the case where price discrimination is prohibited and the incumbent must make the same offer to both buyers. Now, in order to discourage customer entry the incumbent must offer *each* buyer the price p^* . Hence, absent the ability to price discriminate, the monopolist extracts lower rents from customers and discouraging entry is less profitable. Indeed, if p^* is lower than the incumbent's marginal cost, entry deterrence will not occur at the equilibrium and buyers will enter upstream production.

The comparison between price discrimination and price uniformity reveals that price discrimination is always welfare-detrimental in this context. Either price discrimination 'deters' customer entry, or buyers decide to purchase from the incumbent irrespective of price discrimination, but with discrimination one of them pays a higher price thereby producing a deadweight loss.

2.3.2.2 (Explicit and Implicit) Selective Discounts

This section focuses on the case in which the dominant firm and the rival make *simultaneous* offers to *a number of buyers*. In this environment, the dominant firm can make use of selective discounts, whereby it offers a low price to *some key* buyers, while recovering losses on another group of buyers which will pay a high price. In this way, the dominant firm can limit the

²² The price p^* may be above or below cost depending on whether self-supply is more or less efficient than the incumbent's production. See Section 2.4.1.

rival's ability to compete successfully in the market, the more so the more individualised the discounts.

Karlinger and Motta (2012) shows how such a mechanism works in a model characterised by network effects. Technical Section 2.4.2 presents similar results in a model where scale economies are on the supply-side, for consistency with the analysis carried out in the previous chapter. Let us intuitively describe these results. To understand the effects of price discrimination, we first describe what happens when price discrimination is not allowed, and then turn to the case of discrimination.

Benchmark regime: uniform pricing Imagine that there are two buyers (possibly asymmetric), and that the rival firm is more efficient than the incumbent when it supplies both of them, but that the rival is poorly efficient when it serves a single buyer. Suppose also that both the rival and the incumbent make simultaneous price offers to the two buyers.

As also discussed in Chapter 1, such a cost structure can arise in situations in which the rival is an entrant firm that can produce at lower marginal costs than the incumbent, but that has not paid the entry cost yet when competition for buyers starts. Also, selling to a single buyer is not profitable enough to cover the entry cost. We will refer to this case in the more formal analysis of Section 2.4.2. However, the mechanisms highlighted in this section also apply to situations in which *the rival is already in the market* and the incumbent preserves an advantage in the supply of a single contestable buyer because of the existence of scale economies, learning effects or network externalities that the rival has not fully exploited yet.²³

In this context, if price discrimination is forbidden, it can be shown that either of two outcomes is always possible: (i) There exists an 'entry/expansion equilibrium' where both buyers buy from the more efficient rival, paying a price which equals the marginal cost of the incumbent. But (ii) there also exists an 'exclusionary equilibrium' where both buyers buy from the incumbent, at a price which may even be as high as the monopoly price.²⁴

The second result may appear very puzzling at first sight to the reader. How is it possible that the incumbent, who is less efficient than the rival

²³ For instance, in Karlinger and Motta (2012) scale economies are on the demand side, and in an exclusionary equilibrium the rival – which is already in the market but has a weaker installed base of buyers – will not manage to serve new buyers.

²⁴ The model with simultaneous and uniform price offers is also discussed and analysed in Chapter 1 (see, specifically, Sections 1.2.3.2 and Section 1.3.2.2, for the case of symmetric buyers).

in supplying both buyers, manages to serve both of them, and for good measure also at a price which may be well above its marginal costs? The reason is the existence of scale economies and the fact that serving one buyer is not sufficient to recover the entry cost (or more generally to achieve efficient scale). This introduces an externality among buyers. If one buyer chooses to buy from the incumbent, the other buyer knows that she alone would not be sufficient to guarantee profitability to the rival. Even if the rival offers a lower price than the incumbent, there is no point in accepting the rival's offer, as the buyer will end up buying from the incumbent as well. Then, she will have no other option than accepting the incumbent's offer. This creates the scope for *coordination failures*: if each buyer expects the other to buy from the incumbent, both of them will end up buying from it, even though they would have been better off by both buying from the rival.

When there is a multiplicity of equilibria, like in this case, experimental evidence can provide some guidance to establish which type of equilibrium is more likely to arise. The insights offered by Landeo and Spier (2009) and Boone et al. (2014) are particularly interesting for us. Even if in their framework the incumbent uses exclusive contracts to exclude, the nature of the coordination problem is the same as ours. These studies find that exclusion due to coordination failures occurs surprisingly often. For instance, Landeo and Spier (2009) show that – absent communication among buyers – the exclusionary equilibrium is played by laboratory agents 92 per cent of the time, whereas communication reduces it to 42 per cent. Also, the surplus obtained by the buyers sharply increases with communication.²⁵ This suggests that entry/expansion may be facilitated not only by central purchasing agencies (which group orders of affiliated buyers and create sufficient buyer concentration) but also by looser institutional settings which favour communication among buyers.

Explicit price discrimination Consider now the case where sellers can discriminate prices between buyers. Under discrimination, the same mis-coordination logic described in the case of uniform prices explains why the exclusionary equilibrium would still exist: even if the rival sets a price which is below the one of the incumbent, if one buyer expects the other to buy from the incumbent, then she will do so as well. A possible outcome is therefore the exclusionary equilibrium, at which buyers may pay up to the monopoly price.

²⁵ We discuss these findings more extensively in Chapter 3 (on exclusive contracts).

However, the entry/expansion equilibrium does not necessarily exist any longer under price discrimination. Consider the natural equilibrium candidate, which is the one where the rival (firm E) sets a price equal to (or a shade below) the incumbent's marginal cost $p_I = p_E = c_I$ and both buyers buy from firm E . Given that firm E sets $p_E = c_I$ to both buyers, firm I could obtain all the demand and make positive profits by offering a price slightly below c_I to one buyer and the monopoly price p_I^m to the other. It is straightforward that the first buyer would prefer to buy from I than from E in this case. But also the second buyer will buy from I , even at the monopoly price. The reason is that the second buyer knows that selling to one buyer at a price $p = c_I$ is not sufficient for the entrant to cover the fixed cost (or more generally, for the rival to cover the costs to serve a single buyer). Then, she anticipates that entry/expansion will not occur and that she will end up paying the monopoly price anyway. This shows that price discrimination allows the incumbent to exploit the externality that one buyer exerts on the other and break entry/expansion equilibria more easily as compared to the case of price uniformity.

To find the equilibrium, therefore, we have to identify the prices that the rival must offer to *both* buyers so as to prevent the incumbent from playing this 'divide-and-conquer' deviation, and we have to check whether such low prices are profitable for the rival. Whether this is the case or not, depends on how much more efficient the rival is than the incumbent. If the efficiency gap between the two firms is small enough, then the entry/expansion equilibrium will not exist, and the only equilibrium of the game is the exclusionary one. Therefore, with respect to the regime of uniform pricing, price discrimination reduces the set of achievable (socially efficient) entry/expansion equilibria. In other words, under price discrimination it is more likely that the incumbent will supply the buyers, and at a price which may be as high as the monopoly price, with clear detrimental welfare effects.

Note that when entry/expansion equilibria do not exist any longer, one would not necessarily observe price discrimination at the equilibrium: coordination failures may well sustain an equilibrium in which both buyers are served by the incumbent at the monopoly price. However, it is precisely price discrimination that explains why possible entry/expansion equilibria are broken, by allowing the incumbent to profitably deviate and offer a below-cost price to one buyer and a much higher price to the other buyer.

Implicit price discrimination and other forms of discrimination Price discrimination reduces the scope for entry/expansion also when it is not explicit, but not to the same extent as explicit discrimination. To see why, consider the case where there are *different* buyers, some willing to buy large quantities and some small quantities. If the firms were able to identify buyers, then the incumbent would break entry/expansion equilibria by discriminating explicitly, for instance by attracting the large buyer with a very low price and recouping losses by selling to the small buyer at the monopoly price (as the small buyer alone is insufficient to trigger entry). The incumbent might wish to be very aggressive, and if it was able to discriminate explicitly, it might want to offer a zero price to the large buyer.

If buyers cannot be identified – or if firms are not allowed to offer individualised discounts – then the incumbent may try to achieve the same result by engaging in implicit discrimination, for instance by offering quantity discounts whereby a customer who buys more will enjoy a lower unit price, a zero price in the example we are referring to. However, this scheme would not be successful: at a zero price, both a large and a small buyer would buy; indeed, a small buyer would simply buy a large number of units at zero price, and throw away those she does not need. Consequently, the incumbent will be obliged to raise the price it offers to the large buyer: this will discourage the small buyer from mimicking the large one (the extra units she does not need have to be paid and this is costly), but at the same time this limits the aggressiveness of the price offers. As a result, a quantity discount (implicit discrimination in general) will still reduce the set of parameters for which entry/expansion is possible, but not to the same extent as explicit discrimination. This shows that individualised discounts should raise more concerns than standardised discounts, as their exclusionary effects are stronger, giving some support to the fact that courts and antitrust agencies tend to regard individualised rebates by dominant firms with more suspicion than standardised rebates and quantity discounts.

Similar considerations may apply to other forms of price discrimination. For instance, firms may resort to coupons, which may allow to discriminate not only among consumers (those who receive coupons have the chance to pay a lower unit price) but also among different units sold to the same consumer (a coupon may entitle the buyer who redeems it to a price reduction on one or a limited number of units). In principle, the incumbent may issue a number of coupons that correspond exactly to the number of units that it needs to secure to prevent the rival from reaching the critical size, and recoup the losses caused by the coupons

by selling all the remaining units at a monopoly price. In this way, the incumbent may be able to replicate the same result as under explicit discrimination. However, to the extent that the firms cannot target buyers perfectly, coupons may reach consumers who would not need them. For instance, if firms used random coupons, a small buyer may receive more than the units she wants to consume, and throw away the remaining ones (in consumer markets, it is unrealistic to think that consumers may trade coupons). This would oblige firms to send around more coupons than optimal, which is costly, and will make the discrimination strategy through coupons less aggressive. Like with quantity discounts, coupons will make it more difficult for entry/expansion to take place, but not to the same extent as explicit discrimination.

Welfare considerations and prohibition of below-cost pricing In Section 2.5, we shall discuss the policy implications of the formal economic model, but for the time being it is important to note that not only a ban on price discrimination, but also a prohibition of below-cost pricing would restore the same situation as under uniform pricing (where, however, recall that exclusion may occur because of mis-coordination). Indeed, entry (or, more generally, the rival's expansion) can be impeded because the incumbent plays a 'divide-and-conquer' strategy whereby it makes aggressive offers to one buyer (who pays a price below the incumbent's marginal cost) while recouping losses on the other buyer. If there is a policy rule which prevents the incumbent from setting prices below marginal costs on *any* buyer, then the incumbent will not be able to 'deviate' from entry/expansion equilibria such as the one where the rival sets price equals to the incumbent's marginal cost. Note – in particular – the stress on preventing below-cost pricing on any buyer, as opposed to a rule which simply forbids overall losses: the 'divide-and-conquer' strategy identified above is profitable for the incumbent on average, as it can recover the losses made on some buyers by setting high prices on the other buyers (who will be captive to the incumbent once some key buyers decide not to buy from the rival). This is crucial when discussing policy implications in practice: authorities should not assume a lack of anti-competitive effects just because they observe that on average price discrimination did not entail losses.

Further, the welfare effects of prohibiting price discrimination (or below-cost pricing) are ambiguous. Such a prohibition makes entry/expansion equilibria more likely, but when entry/expansion equilibria exist both with and without price discrimination, imposing

uniform pricing (or prohibiting below-cost pricing) has a chilling competition effect and makes consumers pay a higher price.

Buyer power and buyer coordination As in the model of predation analysed in Chapter 1, the exclusionary mechanism at play here depends crucially on the buyers being unconcentrated and uncoordinated. Buyer fragmentation is crucial because by assumption there is no buyer who single-handedly could command a sufficient volume of business for the rival to be profitable. In our example, this amounts to assuming that neither of two buyers is sufficiently large. In general, this assumption translates into having sufficient fragmentation: as one buyer becomes larger there is a point at which the volume of orders she brings to the rival is sufficient for her orders to lead to profitable entry or, more generally, to achieve efficient scale.

Buyer mis-coordination is also crucial because if buyers could cooperate and jointly decide their purchases, they would internalise the externality that they impose onto each other, and would direct their orders to the rival, which would be better from their collective point of view. This implies that central purchasing agencies would play a pro-competitive role in such circumstances.

Network externalities As already mentioned at the beginning of this section, the mechanism discussed so far applies also to a situation where there are network effects and a critical mass of users needs to be achieved for consumers to derive any utility from sponsoring a given network. In both cases, a buyer deciding to buy from the incumbent exerts a negative externality on the other buyers by making it less likely that the rival achieves its minimum efficient scale. Karlinger and Motta (2012) study extensively the model with network effects and demonstrate the exclusionary potential of explicit and implicit price discrimination. Under uniform pricing both 'entry/expansion equilibria' and 'exclusionary equilibria' exist, whereas under price discrimination only exclusionary equilibria will exist if the incumbent does not have a very strong marginal cost disadvantage vis-à-vis the rival.

Interoperability In situations where there are network effects, incompatibility between networks – that is, the fact that the incumbent's pre-existing customer base does not exert any positive externality on the users of the rival's network - facilitates exclusion. *Vice versa*, a policy rule requiring interoperability between the incumbent's and the rival's network would solve the buyers' coordination problem at the root, making exclusion

impossible in the context of the models we are discussing. Even less than perfect interoperability would work in the same direction by allowing buyers of the rival to enjoy some of the network effect of the incumbent, and would enlarge the set of achievable entry/expansion equilibria. Obviously, interoperability obligations may also have detrimental effects. For instance, they might conflict with intellectual property rights and accordingly have an *ex ante* disincentive effect on investment (see the discussion in Chapter 5). Further, in some cases they may discourage variety and innovations, as the rival might find it more convenient to conform to the specifications of the incumbent's network rather than developing original but incompatible features. However, when there exist strong asymmetries between an incumbent and rivals, interoperability obligations may be a way to avoid those imbalances having long-term consequences (think for instance of imposing roaming obligations on telecom incumbents until new rivals have developed their own network).

2.3.2.3 Discounts conditional on exclusivity

In this section, we summarise the intuition behind economic models that show how discount schemes that include an explicit exclusivity requirement may increase the scope for exclusion. Bernheim and Whinston (1998) is the first paper to deliver this insight. An explicit exclusivity clause allows the incumbent to secure early buyers by offering pricing schemes that entail fewer distortions on the sales made to such critical buyers, thereby earning larger profits. (See Section 2.4.4 for a formal analysis.)

To grasp the main intuition, suppose that (like in the base model of Chapter 1) competition between the incumbent firm and the rival takes place sequentially, in the sense that firms compete first for a cohort of new buyers (or a market) and then for another cohort of buyers (or for another market). Consider the same setting as the one that we have discussed so far, in which scale economies and an incumbency advantage imply that the rival is more efficient than the incumbent in supplying all of the buyers, but it is less efficient in supplying later buyers only. Moreover, imagine that buyer's demand is elastic and that firms compete by offering two-part tariffs, that is, a constant unit price and a fixed fee.²⁶

²⁶ We have considered rigid demand so far because it simplifies the analysis and allowed us to deliver all the important insights. In this case, instead, the assumption of elastic demand is important because it allows us to emphasise the different effects produced by linear and non-linear prices.

In this environment, we know that the incumbent manages to extract more rents than the rival from later buyers (if it secured the early ones), which makes it more aggressive in bidding for early buyers and may lead to inefficient exclusion. As we have already highlighted, the incumbent suffers losses on the early buyers: with two-part tariffs the winning offer entails a unit price equal to the incumbent's marginal cost and a negative fixed fee, that is, a payment from the incumbent to the early buyers. For this reason the loyalty requirement is crucial: without it, early buyers would cash in the fixed payment by purchasing a negligible amount from the incumbent and would buy the rest from the rival. Then, the incumbent would not manage to exclude the rival. As a consequence, should the incumbent be constrained not to include an exclusivity requirement in the offer, it would secure early buyers by setting a below-cost linear price. This introduces an allocative distortion in the sales to early buyers and reduces the incumbent's profits. For this reason inefficient exclusion would be less likely as compared to the case in which the offer can include exclusivity requirements.

Note that what is crucial for the incumbent in order to succeed in an exclusionary strategy is to limit the buyer's purchases from the rival. Under certainty on the buyer's demand, this same outcome may be achieved by making the payment of the compensation conditional on the buyer purchasing a sufficiently high share of its requirements from the incumbent, or on the buyer's purchases exceeding a suitably defined quantity threshold. However, if demand is subject to uncertainty, quantity discounts may be less effective than exclusivity (or market-share) discounts. Similarly, if the suppliers' products are differentiated and buyers value variety, a quantity discount may not be enough to limit buyer's sales from the rival because the buyer, prompted by the low discounted price, would buy a lot from the incumbent, and then may buy enough additional units from the rival.

2.3.2.4 Conditional rebates to target the contestable demand

Another reason why market-share discounts can facilitate exclusion is that they may allow the dominant firm to discriminate the price across captive and contestable portions of the demand, thereby limiting also in this case the amount of profits that need to be sacrificed so as to exclude the more efficient rival, as we show formally in Section 2.4.3.

Consider again the setting of Chapter 1 in which the incumbent and the rival compete for buyers in sequence and assume that the following conditions hold. When competition for the early buyers takes place, the

rival has not achieved efficient scale yet. Suppose also that part of the demand of each buyer is *non-contestable* by the rival; that is, no matter which prices are offered in the market, buyers will satisfy that part of the demand buying from the incumbent. For instance, buyers can be thought of as retailers/firms who serve different categories of consumers, some having a strong preference for the incumbent's product, while others considering the rival's product as a substitute of the incumbent's product. Further, supplying the contestable demand of the later buyers is not enough for the competitor to achieve efficient scale, it also needs to supply the contestable demand of the early buyers.

As already discussed several times, as a consequence of these features the incumbent manages to extract larger rents from later buyers than the rival, which puts it in the position of offering a very large discount to the early buyers, a discount that the rival may be unable to match. Hence, inefficient exclusion arises at the equilibrium if the incumbent's efficiency disadvantage is not too large. The novel insight of this analysis is that exclusion is more likely if the discount offered by the incumbent is *conditional on the buyers' purchases exceeding a suitably defined threshold*. Conditionality allows the incumbent to target the aggressive price offer to the contestable part of the demand. For instance, if each buyer bought 100 units for sure from the incumbent, and would consider the entrant only for the purchase of additional units, the incumbent would offer a discount conditional on buying more than 100 units. If, instead, the incumbent could not discriminate, the incumbent should set a below-cost price also for the captive units demanded by early buyers. Then, the amount of profits that the incumbent should sacrifice so as to exclude would be larger.

Moreover, in a setting in which buyer's demand is certain and common knowledge, what really matters to target the contestable demand is that the buyer qualifies for the discount if her purchases exceed a given threshold. Whether the threshold is expressed in terms of units (quantity discount) or share of total requirement (market-share discount) is irrelevant. If demand, however, is subject to shocks, market-share discounts may be more effective than quantity discounts to target the contestable demand.

Finally, the discount offered to the early buyers entails that the contestable units are effectively sold below costs. However, on the non-contestable units the price charged to the early buyers may be sufficiently high that there is an overall positive profit out of these buyers. In terms of policy implications, this result warns us on how to properly conduct the price-cost test: it is not on all the units sold to a buyer that the

price-cost test should be implemented, but rather on the contestable units of that buyer. Moreover, as shown in the formal analysis of Section 2.4.3, in performing the test one should compare to costs the appropriate discounted price, that is, the *effective* price that the incumbent is setting for the contestable units, which is the actual price that the rival should match so as to capture the contestable demand. To do so, one should attribute to the contestable units the *entire* discount that the buyer loses when she addresses the rival and does not qualify for the discount. When conditional rebates are retroactive, the identification of the effective price requires additional care, as we discuss in the next section.

2.3.2.5 Retroactive rebates

Retroactive rebates are conditional rebates such that, once a target threshold is reached, the buyer receives the discount on *all* the units previously purchased. They are also denoted as ‘all-units’ rebates, or ‘back to \$1’ rebates. The rebate is, instead, incremental, if the discount applies only to the units exceeding the target threshold.

If the rival can contest only part of buyers’ demand, then the discount it will have to offer in order to match a retroactive rebate by the dominant firm is much higher than it would appear by simply looking at the nominal discount offered by the dominant firm.

Suppose that there is a single buyer in a market whose demand is denoted by Q . Suppose that a proportion $s \in (0, 1)$ of this demand is contestable, while the remaining proportion $1 - s$ is captive to the incumbent. Suppose now that the incumbent firm offers a per-unit discount d (in percentage terms) on the list price p_I , *conditional on buying H units from it with $H > (1 - s)Q$* . The discount d applies to all the units bought by the buyer, once the target threshold H is achieved. What is the effective price that rival firm E has to pay in order for the buyer to prefer buying the contestable units from E rather than from the incumbent?

If the buyer buys the contestable units from the incumbent, then she will qualify for the discount thereby spending $p_I(1 - d)Q$ for her total requirement. If, instead, the buyer buys the contestable units from the rival, then she loses the discount on all of the non-contestable $(1 - s)Q$ units that will be bought from the incumbent at a full price. The buyer’s total expenditure will be $p_E sQ + p_I(1 - s)Q$, where $p_E sQ$ is the sum paid to the entrant to buy the sQ contestable units and $p_I(1 - s)Q$ is the sum spent to buy from the incumbent the $(1 - s)Q$ captive units. Hence, the buyer will

prefer to satisfy the contestable demand from the rival if:

$$p_E \leq p_I \left(1 - \frac{d}{s}\right) \equiv p_{eff} \quad (2.1)$$

where p_{eff} is the effective price the rival has to offer in order to match the incumbent's discount. This is the price that should be compared to relevant measure of costs in order to perform a price-cost test. (See the discussion below.)

Note that when all the demand is contestable, that is, $s = 1$, the effective price coincides with the nominal discounted price $p_I(1 - d)$ offered by the incumbent. When, instead, not all the demand is contestable, that is, $s < 1$, the effective price is lower than the nominal discounted price. In fact, the above formula attributes to the contestable units the discount that the buyer loses on the non-contestable units when she addresses the rival (and does not qualify for the discount). In total, the lost discount amounts to $(1 - s)Qdp_I$, which translates into $\frac{(1-s)Qdp_I}{sQ}$ per contestable unit. By adding this amount to the nominal per-unit discount dp_I , one can find the actual price that the rival must set so as to outweigh the retroactive discount of the incumbent:

$$p_I - dp_I - \frac{(1-s)Qdp_I}{sQ} = \left(1 - \frac{d}{s}\right)p_I \equiv p_{eff} \quad (2.2)$$

For instance, if the list price is $p_I = 100$, the discount offered by the incumbent is $d = 10\%$ and the contestable share of demand is $s = 0.25$, then the effective price is:

$$p_{eff} = 100 - \underbrace{10}_{10\%(100)} - \frac{10(0.75)Q}{0.25Q} = 100 - 10 - 30 = 60 \quad (2.3)$$

which is much lower than the nominal discounted price set by the incumbent ($90 = 100 - 10\%(100)$). In other words, in this case, a nominal discount of 10 per cent requires the rival to offer an effective discount of 40 per cent to be able to attract the contestable demand.

The above expression also highlights that the higher the contestable share the higher the effective price, because the total discount lost on the non-contestable units is, *ceteris paribus*, smaller. For instance, if the contestable share increases from $s = 0.25$ to $s = 0.5$, then the effective price increases from 60 to $80 = 100 - 10 - \frac{10(0.5)Q}{0.5Q}$, and the nominal 10 per cent discount lower translates into an effective discount of 20 per cent. From the expression above one can also notice that the effective price

of the *marginal* unit, that is, the one that makes a buyer qualify for the discount, is necessarily lower than the effective price that the rival must offer to secure the entire contestable demand. Indeed, the discount lost on the non-contestable units is attributed to a single unit, rather than to the entire contestable demand, when one computes the effective price for the marginal unit. Consider again the example in which the list price is $p_I = 100$, the discount offered by the incumbent is $d = 10\%$ and the contestable share of demand is $s = 0.25$. Assume that the buyer's demand is $Q = 100$. In this case, the effective price on the 76th unit (the marginal one) turns out to be negative: $p_{eff} = 100 - 10 - \frac{10(0.75)100}{1} = -660!$ This clarifies that, if one wanted to assess the anti-competitive potential of retroactive rebates, it would not be reasonable to focus only on the effect of the rebate scheme on the marginal unit.

We conclude this section with three remarks. First, a retroactive rebate can always be replicated by an incremental rebate. Following the example above, a retroactive rebate of 10 per cent if the buyer buys more than $H = 0.75Q$ units from the incumbent (which corresponds to an effective discount of 40 per cent) would be equivalent to an incremental rebate of 40 per cent if the buyer buys more than $H = 0.75Q$ units from the incumbent. However, it is undeniable that the retroactive rebate 'masks' a strong discount that would be immediately transparent should the firm use an incremental rebate. This is perhaps a reason why they typically raise greater concerns than incremental rebates.

At the same time, one should take into account that retroactive rebates are clearer and simpler to manage: this is probably why they are so popular. So far, we have reasoned as if each buyer makes just a single purchase. But buyers and sellers typically make repeated transactions within a certain period.²⁷ The advantage of retroactive rebate is that the buyer pays the same price for each unit during the reference period and calculations on the amount of money that the seller has to return to the buyer are made at the end of the period. With an incremental rebate, instead, different units sold during the reference period may have a different price, making it complex for buyers and sellers to handle the transactions.

However, suppose the seller is uncertain on how many contestable units the buyer will purchase. Suppose also that it wants to set a certain price

²⁷ Chapsal (2014) studies a model in which scale economies are prevalent and buyers engage in repeated purchases. In a setting in which the incumbent can move first, he shows that conditional discounts can allow the incumbent to exclude a more efficient rival that enters in a later period by committing to a very low price on the units demanded by the recurrent buyers in later periods.

on the contestable units. An incremental scheme allows the seller to set that price with certainty, no matter how many contestable units the buyer will demand. Instead, as we discussed above, in the case of a retroactive rebate the effective price for the contestable units depends on the size of the contestable demand. If the buyer ends up demanding very few contestable units, the effective price may be very low and the retroactive rebate may turn out to be more costly for the seller than the incremental one.

Application of the price-cost test to retroactive rebates In cases of allegations of price-based exclusionary conduct, a competition authority will typically want to perform a price-cost test to establish whether the price charged by the dominant firm is above or below a relevant measure of costs. In case the conduct under examination is a retroactive rebate, the relevant price the agency should look at will be the ‘effective’ price that an as-efficient rival will have to charge to be able to win the contestable share of the market. Such effective price is precisely the one derived in equation (2.2).²⁸

When carrying out such a test, one must take into account that the calculation of the effective price is *very sensitive* to the estimates of the contestable share of demand, especially at low levels of the contestable share (that is, when s is small, a change in the contestable share has a big impact on the estimated effective price and thus on the chance that this may be above the relevant cost measure).

Furthermore, the measurement of the contestable share also affects the estimation of the cost benchmark the effective price should be compared to. Since fixed costs are included in total costs (and in long-run incremental costs) and are partially included in avoidable costs and short-run incremental costs (see the discussion in Chapter 1), the average level of these costs depend on the total contestable demand over which they are computed. Hence, a higher contestable share leads not only to a higher effective price but also to a lower cost benchmark, thereby making it less likely that the discounted price turns out to be below cost.²⁹

Since it is quite difficult in practice to assess which share of the demand is contestable and which share would instead stick to the dominant firm’s product *no matter* the price difference with competing products, the above considerations suggest that this price-cost test should be implemented with care, and complemented by a serious sensitivity analysis which

²⁸ The price-cost test applied in the context of retroactive rebates commonly referred to as the ‘as-efficient competitor test’ (see also the discussion in European Commission, 2009).

²⁹ Due to the existence of scale economies, the cost benchmarks should be computed aggregating the contestable demand of the customers that are offered a rebate scheme.

demonstrates that the conclusions obtained do not depend on a specific value of the contestable share.

Another issue that concerns the measurement of the contestable share relates to the time horizon over which one should assume that a buyer is able to switch some of its purchases away from the dominant firm. The longer this time horizon, the higher the contestable share. Note that the identification of the appropriate time horizon is also related to the nature of the buyers targeted by the discount. As stressed by Federico (2011a) if there are crucial buyers whose purchases could legitimise an entrant, thereby leading to follow-on sales with other buyers, it would be appropriate to use a longer time horizon in the computation of the contestable share.

2.3.3 Other Anti-competitive Concerns of Contracts that Reference Rivals

Some recent works have provided additional arguments in support of the view that loyalty rebates – or more generally contracts that reference rivals (such as market-share discounts offered by a seller, for example, which also depend on how much a buyer purchases from that seller's rivals) – raise more severe anti-competitive concerns than contracts that *do not* condition the terms of trade on how much the buyer purchases from rivals. We next review some of this academic economic literature.

2.3.3.1 Exclusivity discounts

Calzolari and Denicolò (2013, 2015) and Calzolari et al. (2016) consider a *one-period* model of price competition, in which the distinct feature of the dominant firm is not the incumbency advantage, as we have assumed so far; rather dominance stems from a competitive advantage vis-à-vis the rival, in the form of higher quality of the product or lower cost.³⁰ Moreover, the contracts whose effects are explored in those papers are denoted as

³⁰ The standard assumption in the literature on exclusion is that the rival is more efficient than the incumbent (the dominant firm). Under level-playing-field competition, one would expect that the more efficient rival manages to operate successfully in the market. Then, if exclusion occurs, it must be that other asymmetries between the dominant firm and the rival produce an advantage for the former. The purpose of the literature is to identify such asymmetries (for instance, an incumbency advantage, a strategic advantage, etc.) and the mechanisms through which they allow the dominant firm to exclude the more efficient rival. In the papers we are discussing in this section, the dominant firm has superior quality or uses a superior technology. In a sense, dominance is benign. None the less, these papers show that exclusivity rebates can be used by the dominant firm for anti-competitive purposes.

exclusive dealing contracts, but they amount to what we define in this book as loyalty or exclusivity discounts: they do not involve a (long-term) commitment on the side of the buyer; it is only the supplier that commits to offer terms of trade that vary depending on how much the buyer purchases.

These papers rely on a mechanism which is different from the one highlighted in this chapter (and in this book), based on the idea that loyalty rebates (and other practices) may be used by the dominant firm to deprive the rival of the crucial scale it needs to be viable. The key reason why exclusivity rebates may turn out to be anti-competitive is *imperfect rent extraction*, namely the fact that the dominant firm cannot fully extract the buyers' surplus by means of non-linear contracts, for instance through fixed fees. Imperfect rent extraction may arise because the buyers' surplus is private information and suppliers do not know exactly how large it is, or it may be due to buyers being risk- (or loss-) averse.

Because of imperfect rent extraction one departs from the so-called *neutrality result* due to Bernheim and Whinston (1998). According to that result, under complete information and non-linear pricing, exclusivity discounts (or market-share discounts) are irrelevant. The intuition is that, by means of non-linear prices, a firm can extract all of the buyer's surplus in excess of what the buyer obtains by trading with the competitor. As a consequence, each supplier offers a contract that maximises that surplus – for instance by pricing at marginal cost – and extracts all of it through the fixed fee. Absent inefficiencies in contracting, it follows that equilibrium contracts maximise *total surplus* from trade. Therefore, if trading with a single supplier (say the dominant firm) is inefficient – for instance because suppliers offer differentiated products and trading only with the dominant firm reduces total surplus – then the equilibrium outcome will not involve trade with a single supplier. If, instead, trading with a single supplier maximises total surplus, then it will arise as the equilibrium outcome but as a unilateral decision of the buyer, without the need to introduce an explicit requirement in the contract that conditions the terms of trade to the whether exclusivity is satisfied. In other words, exclusivity discounts (or market-share discounts) are either unprofitable or superfluous. For this reason, they should not raise the concern of competition authorities.

Two routes have been explored to overcome the neutrality result. One is to consider imperfect rent extraction, as in Calzolari and Denicolò (2013 and 2015). The other is to maintain the assumption of complete information and introduce inefficiencies in contracting, as in Bernheim and Whinston (1998) and in the literature that originated from that paper.

Exclusivity Discounts and Imperfect Rent Extraction To see why contractual exclusivity may not be superfluous, imagine that the buyers' willingness to pay is private information. Then suppliers find it optimal to introduce a distortion in the contract designed for buyers with low willingness to pay so as to extract more surplus from buyers with high willingness to pay while inducing them to reveal their type. When contracts cannot reference rivals, the distortion consists of reducing sales to low-type buyers below the efficient level, for instance by offering a two-part tariff which involves a variable price above costs. Such a distortion facilitates rent extraction from high-demand buyers because they would lose a considerable amount of surplus if they mimicked low demand buyers and had to reduce their purchases. However, the distortion reduces the supplier's profits. A more profitable way to extract surplus from the high-valuation buyers is to deal in exclusivity with low-demand buyers. The model assumes that buyers have a preference for variety. Hence, a high-demand buyer would suffer too high a loss if she tried to mimic a low demand buyer and was obliged to purchase from a single supplier. Then, the introduction of contractual exclusivity facilitates rent extraction from high-demand buyers with the advantage that the distortion reduces the *rival's sales* to low-demand buyers, *not own sales*. Indeed, own sales to low-demand buyers increase under exclusivity, if suppliers offer (imperfect) substitute products. This gives the incentive to the suppliers to create even more distortions by involving a larger set of buyers in exclusivity.

Next, let us add competition to the picture. When the asymmetry between the dominant firm and the rival is high, because the quality/efficiency gap between them is large, the dominant firm does not suffer much competitive pressure from the rival. Then, the dominant firm must not concede a large discount to make a buyer purchase in exclusivity. In this case, contractual exclusivity (or more generally market-share discounts) benefits the dominant firm but harms total welfare – as compared to the case in which contracts cannot reference rivals – because the buyer suffers from a larger distortion, and such a distortion is absorbed by the rival whose access to low-demand buyers is foreclosed. When, instead, the asymmetry between the dominant firm and the rival is limited, the introduction of exclusivity intensifies competition. Absent exclusivity, firms compete for marginal units of a buyer, and the presence of product differentiation makes competition softer. Instead, with exclusivity firms compete for the entire requirement of a buyer. The presence of some differentiation between the suppliers' products does not matter for the outcome of competition, what matters is the amount of rents that each

supplier is able to leave to that buyer. This makes competition tougher, and equilibrium prices and profits lower, as compared to the case in which contracts do not reference rivals. Then, exclusivity discounts exert a pro-competitive effect when asymmetry between suppliers is limited.³¹

The mechanism that we have just discussed applies not only to situations in which buyers have private information. The crucial factor for the argument to apply is that, for some reason, sellers cannot extract the entire buyers' surplus through a fixed fee. Imagine, for instance, that there exists demand uncertainty (see Calzolari et al., 2016). Under demand uncertainty, fixed fees may entail a loss for the buyer. Then, if buyers are loss-averse, fixed fees can extract only limited rents from buyers. To facilitate rent extraction, suppliers need to introduce a distortion, that is, they have to increase the marginal price above costs. Furthermore, contractual exclusivity makes the distortion more profitable, because it increases own sales when suppliers' products are (imperfect) substitutes. When the asymmetry between the dominant firm and the rival is pronounced, contractual exclusivity is profitable for the dominant firm and welfare-detrimental. When the asymmetry is limited, contractual exclusivity turns out instead to be pro-competitive.

Contracting Externalities and Contractual Exclusivity Contracting externalities exist when the payoff of a contracting agent depends not only on her own terms of trade but also on the terms of trade obtained by other contracting agents.

A setting in which contracting externalities arise is the model with scale economies and sequential buyers analysed in Chapter 1, dealing with predation, and that we have analysed in this chapter to study exclusivity discounts and market-share discounts. (We will refer to variants of that model also in the next chapters of the book.) In that setting, the payoff of the incumbent depends not only on how much it sells to the first buyer, but also on how much the rival supplier sells to the first buyer. The reason is that how much the rival supplier sells in the first period determines its second period cost and then the rents that the incumbent is able to extract from the second buyer.

As we have already discussed, in that context the more efficient rival can be excluded from sales in period 1 because exclusion, by softening

³¹ In this case, firms' profits are higher absent exclusivity. However, each firm has a unilateral incentive to introduce it. As a consequence, they end up both offering exclusivity and earning lower profits.

competition among suppliers in period 2, allows the incumbent to extract more rents from the second buyer, that is, from the agent not involved in the initial negotiation.³² Furthermore, the presence of contracting externalities explains why *contractual exclusivity is not superfluous*: indeed it favours exclusion, as we have discussed in Section 2.3.2.3, by allowing the incumbent to implement more profitable non-linear pricing schemes.

Contractual externalities also naturally arise in a setting in which buyers are retailers/firms that compete in a downstream market and the profitability of selling in that market depends on the own terms of trade for the input but also on the input price paid by rivals. In that environment, the joint surplus of the contracting agents may be maximised when trade occurs with all the suppliers and/or with all the retailers, but it may happen that such a maximal joint surplus cannot be sustained at the equilibrium if contracts are not rich enough to internalise all the externalities that arise in the negotiations. Then, even though trading with a single supplier (or with a single retailer) is inefficient, because it reduces total surplus relative to the maximal amount that can be achieved, the equilibrium outcome may entail trade with a single supplier (or with a single retailer) because it is the second-best solution that contracting agents are able to sustain.

An illustration of this idea is given by Hart and Tirole (1990).³³ They show that a monopolist supplier, trading with multiple retailers, may be unable to sustain the outcome that maximises the joint industry profits (that is, the vertically integrated outcome). The reason is that when offers are private, that is, when the terms of the contract that the upstream monopolist offers to each downstream firm cannot be observed by the other downstream rivals, then the monopolist suffers from opportunistic behaviour.³⁴

³² Exclusion of the rival arises at the equilibrium because it maximises the joint payoff of the agents involved in the initial negotiation. However, since the payoff of the second buyer is not internalised by the agents negotiating initially, exclusion turns out to be detrimental to total welfare. In the terminology of Bernheim and Whinston (1998), this is a setting with ‘non-coincident markets effects’. See also Spector (2011), who studies the relative importance of the absence from the contracting game of some affected parties and of the restrictions imposed on the set of feasible contracts for inefficient outcomes to arise.

³³ See also the subsequent work by O’Brien and Shaffer (1992), McAfee and Schwartz (1994) and Rey and Vergé (2004). See also Rey and Tirole (2007) for an insightful review of this literature.

³⁴ Equivalently, one may assume that if such a contract was publicly observable, it could also be privately renegotiated; that is, it is impossible for the upstream monopolist to credibly and publicly commit to a certain price for the sale of the input.

Let us explore this mechanism through a stylised example. Assume that there are two independent downstream firms, D_1 and D_2 , which sell a (imperfectly) substitutable final product. Suppose that the retail prices p_1^* and p_2^* with the associated sales q_1^* and q_2^* are the ones that a firm, vertically integrated with both downstream units, would choose so as to maximise industry profits. Imagine also that there exists a wholesale price w^* such that, if D_1 and D_2 paid w^* for each unit of the input, then they would set p_1^* and p_2^* in the final market, selling q_1^* and q_2^* and making profits equal to $\pi_1^* = (p_1^* - w^*)q_1^*$ and $\pi_2^* = (p_2^* - w^*)q_2^*$. The upstream monopolist U_I could then extract all the vertically integrated profits in the market if it was able to convince D_1 and D_2 to accept a contract whereby they buy each unit of the input at a price w^* and in addition pay a fixed fee π_i^* to U_I .³⁵ But the downstream firms will not want to sign such a contract. To understand why, suppose that they did sign such a contract and consider what would happen next.

After agreeing on the above terms, with each downstream firm having paid π_i^* to U_I , the upstream monopolist would have an incentive to renegotiate and offer either firm, say D_1 , the input at a slightly lower unit price than w^* . This would allow D_1 to have a lower input cost than its rival, sell a quantity $q' > q_1^*$ (if competition was very fierce, it might even serve the whole market), and earn profit $\pi' > \pi_1^*$. Therefore, D_1 would be willing to pay up to π' for the new contract, giving the incumbent an additional profit $(\pi' - \pi_1^*)$.

Note, however, that firm D_2 's profits would fall as a result of this, since the original contract commits it to pay π_2^* , but after the renegotiation between U_I and D_1 , it would sell less and earn less than π_2^* (if competition was very fierce, it would sell and earn nothing). Since D_2 will anticipate the upstream monopolist's temptation to renegotiate the contract with D_1 , it would then be unwilling to sign a contract with the upstream monopolist under which it pays π_2^* . Note that the same might happen with either of the two downstream sellers, so neither would be willing

³⁵ The fact that at this contract they would make zero profits is *not* what will make them reject the contract. We are assuming that the upstream monopolist has all the bargaining power and that if they rejected the offer, the downstream firms would not have the input, and would therefore make zero profits. So if they did make zero profits, they would still accept the contract. (If the reader is uncomfortable with the idea that the gain is exactly zero, one can posit that the upstream monopolist asks for a fixed payment which is slightly less than π_i^* , so that D_1 and D_2 would make strictly positive profits.) The problem comes from the fact that if they accepted such contracts they would make negative profits, as explained below.

to enter into an agreement with U_I unless a very low fixed payment is set.

The conclusion of this literature is that, if the monopolist supplier trades with both retailers, the contracts that are accepted at the equilibrium sustain joint profits that are much lower than the vertically integrated ones. Then, it is in the joint interest of the contracting parties that the supplier distributes its product through a single retailer, committing not to trade with the rival downstream firm. Under product differentiation in the downstream market, trading with a single retailer makes profits lower than the vertically integrated ones. However, trading with a single retailer may represent a second-best outcome that gives larger profits to the contracting agents than those obtained when trade occurs with both retailers. Note that the equilibrium contract needs to impose an explicit exclusivity requirement – on the side of the upstream monopolist that commits not to trade with other downstream firms – so as to remove the opportunistic behaviour. Hence, *contractual exclusivity is not superfluous* in this setting. An alternative solution for the monopolist supplier so as to avoid opportunism is to integrate vertically with a downstream firm and to refuse to supply the independent downstream firm. We will discuss this issue when we discuss vertical foreclosure in Chapter 5.^{36,37}

2.3.3.2 Market-share discounts and buyers' opportunism

Choné and Linnemer (2014) consider a setting in which the incumbent and a buyer negotiate on non-linear pricing schemes at a time in which

³⁶ Other possible ways to solve the problem include resorting to resale price maintenance, to most favoured nation clauses, or simply by reputation: if U_I and the downstream firms were going to interact repeatedly over a long horizon (as may presumably be the case in many supplier-distributor relationships), then the upstream monopolist may be able to solve the commitment problem simply by establishing the reputation of not renegotiating contract terms.

³⁷ Marx and Shaffer (2007) show that, when contractual offers are (simultaneously) made by downstream firms, then the existence of contracting externalities prevents the agents from sustaining the vertically integrated outcome, even though the supplier's acceptance decisions are public and firms can use three-part tariffs. (Three-part tariffs are contracts including a fixed upfront payment, paid when the contract is signed, and another non-linear component paid if and only if a positive input quantity is purchased.) Also in this case, the second-best outcome where the supplier trades with a single retailer arises. However, if retailers can offer contracts that are contingent on whether trade occurs in exclusivity (see Miklos-Thal et al., 2011) or if they can offer a menu of contracts (see Rey and Whinston, 2013), then the vertically integrated outcome can be sustained. These results emphasise that sufficiently rich contracts manage to internalise all the externalities that arise in contracting and to avoid inefficient outcomes (for the vertical structure).

the characteristics of the rival, for instance its marginal cost or the quality of its product, are not yet known. The purpose of the rebate scheme is to place the buyer in a favourable position when bargaining with the rival and obtain a price low enough to compensate for the discount lost when she purchases less from the incumbent. In other words, the incumbent-buyer coalition uses the rebate scheme to extract rents from the rival.³⁸ However, rebate schemes may also create the scope for opportunistic behaviour of the buyer, who *ex post* has an incentive to purchase inefficiently many units from the incumbent to pocket the discount. The paper compares rebate schemes that reference the rival and rebate schemes that do not: the former allow the incumbent to eliminate buyer opportunism, but the rival's supply is distorted more downwards as compared to non-conditional rebates. This explains why conditional rebates may turn out to be more detrimental for welfare than non-conditional rebates.

2.3.4 Price Discrimination between Off-net and On-net Users

In this section we discuss, with reference to some academic economic literature, some price discrimination issues that arise specifically in the telecommunications sector. Telecommunication networks need access to customers attached to rival networks in order to enable own users to communicate as extensively as possible. A network terminating a call on the rival network must typically pay an access charge (or termination rate) to the receiving network. A concern that has been often expressed is that incumbent networks may be tempted to charge prohibitively high termination charges to place smaller networks at a competitive disadvantage and to foreclose their activity.³⁹

Lopez and Rey (2016) rationalise this concern. The mechanism is the following. Above-cost termination charges create *price-mediated network externalities*: off-net calls (that is, those between different networks) become more expensive than on-net calls (that is, those between the same network) which makes users prefer a large network, for which a higher proportion of calls remain on-net. When termination charges are moderate, network externalities are not strong. This feature, together with

³⁸ The role of rebate schemes is similar to the one played by penalties for breach of exclusivity in the model of Aghion and Bolton (1987) that we discuss in Chapter 3.

³⁹ See for example European Regulators Group (2008); Autorité de régulation des communications électroniques et des postes (French telecommunications regulator), Decision 2007-0810, 4 October 2007; and Comisión del Mercado de las Telecomunicaciones (Spanish telecommunications regulator), Decision AEM 2006/726, 28 September 2006.

a moderate incumbency advantage – the paper assumes that all users are initially attached to the incumbent’s network and incur a switching cost to move to the entrant’s network – implies that competition for users results in the market being shared by the two networks. However, when access charges are high, network externalities become important, which makes users prefer to join the same network, either the incumbent’s or the entrant’s. In this environment, competition for users may result in everybody choosing the incumbent’s network. This happens when users suffer from inertia (that is, when, in the presence of multiple equilibria in the users’ choice, the equilibrium favourable to the incumbent always realises) or when the incumbency advantage is sufficiently strong (that is, the switching cost is large enough). In these cases it is profitable for the incumbent to choose sufficiently high access charges, thereby foreclosing the entrant’s activity.^{40,41}

It is important to comment on the interpretation of the above result. Lopez and Rey (2016) demonstrate that exclusion through the manipulation of termination charges is possible, but particular conditions need to be satisfied for this result to be valid. First, as mentioned above, it is necessary either that the incumbency advantage is strong or that consumers suffer from coordination problems. Second, choosing high termination charges is profitable for the incumbent only when it allows to foreclose the entrant’s activity entirely. This limits the exclusionary concern to markets where potential entrants have been unable to build any customer base. Third, foreclosure is no longer possible under the ‘receiver pays principle’ regime, because all usage prices (even on-net

⁴⁰ This result contrasts with the one obtained by a different literature whose general insight is that competing networks, when they are not too asymmetric, would rather opt for low termination rates, even below costs, than for high termination rates. This conclusion was reached by Dessein (2003), accounting for demand expansion effects, and by Laffont et al. (1998b) and Gans and King (2001). These latter papers allow, as in Lopez and Rey (2016), for off-net/on-net price discrimination. In this environment, below-cost termination charges, by making off-net calls cheaper than on-net calls, make users prefer small networks. This will soften competition for users, thereby allowing networks to sustain high subscription fees and to make larger profits than in the case where termination charges are above costs. Lopez and Rey (2016) shows that preference for below-cost termination charges is valid as long as one restricts network asymmetries and termination charges to take moderate values.

⁴¹ The literature on the choice of the termination charges was started by Laffont et al. (1998a) and Armstrong (1998). These papers show that high termination charges may act as a collusive device allowing networks to sustain high retail prices and profits. Already in the same article, Laffont et al. emphasise that the previous result is not valid if networks offer two-part tariffs instead of simple linear usage prices.

ones) are then set at the off-net cost. Finally, the use of termination charges to exclude may be difficult when one accounts for firms' heterogeneity in costs and quality and for the possibility that entrant networks design a commercial strategy so as to target specific categories of users.⁴²

2.4 Price Discrimination when Scale Matters*

2.4.1 Innes and Sexton's 'Divide-and-Conquer' Price Discrimination*

In a market there exist two perfectly identical buyers, B_1 and B_2 , each with demand $q = 1 - p$. A monopolistic incumbent firm, I , has already sunk its costs, and produces the good at a constant marginal cost $c_I < 1/2$. There are no competing suppliers that can enter this market, but the buyers may consider 'integrating upstream' (alone or together) and engage in self-supply. To do so, they should incur the fixed (set-up and organisation) cost f , and they would then produce at zero marginal cost.⁴³ Like Innes and Sexton (1993), we ignore issues of coordination in self-supply and we simply assume that they will be able to establish production and share the good whenever jointly convenient. Like them, we also assume that:

$$\underline{f} \equiv \frac{1}{2} - \frac{(1 - c_I)^2}{8} < f < 1 - \frac{(1 - c_I)^2}{4} \equiv \bar{f}, \quad (2.4)$$

where the second inequality implies that if the incumbent tried to impose the monopoly price to both firms, they would find it profitable to set up a joint venture to self-supply the product, while the first inequality implies that a buyer will never find it convenient to establish self-supply by itself.

⁴² Calzada and Valletti (2008) provide another justification for incumbent networks to favour high termination charges because of their effect on entry patterns. For a given number of firms in the industry, higher termination charges decrease individual profits. However, when new operators face entry costs, higher termination charges also limit entry. The net effect is an increase in the profits of incumbent networks. However, this result crucially relies on the incumbents' ability to commit not to decrease the termination charge if entry takes place.

⁴³ Innes and Sexton (1993) assumes that the marginal cost involved in self-supply is the same as the incumbent's. We assume here that the marginal cost of self-supply is lower than the incumbent's so as to obtain a richer set of results. See the discussion below.

We shall contrast the case of *uniform pricing*, where the incumbent cannot discriminate between buyers, with the case of *price discrimination*, where it can offer different prices to them.

The timing of the game is as follows:

0. Firm I publicly announces prices p_1 and p_2 to buyers B_1 and B_2 , and commits to them. (Under uniform pricing, $p_1 = p_2 = p$.)
1. Buyer B_1 decides whether to accept or reject the offer.
2. Buyer B_2 decides whether to accept or reject the offer.
3. Buyers who have accepted the offer pay their order and consume. Buyer(s) who have rejected it, can pay f and produce the good (and consume it) at cost 0.

Buyers' decisions As usual we move backwards, and consider buyers' choices first, abstracting on whether prices can be different or not.

At stage 2, B_2 has to decide whether to accept price p_2 or engage in self-supply (possibly jointly with B_1). If B_1 had accepted the offer, then B_2 would accept any offer involving $p_2 \leq p^m$ where $p^m = (1 + c_I)/2$ is the incumbent's monopoly price. To see why, note that the buyer will prefer buying from I as long as the surplus it derives from the purchase is higher than the surplus it derives from self-supplying alone (given that B_1 buys from I): $CS(p_2) \geq CS(0) - f$, that is:

$$\frac{(1 - p_2)^2}{2} \geq \frac{1}{2} - f.$$

For $p_2 = p^m$, the previous inequality becomes:

$$\frac{(1 - c_I)^2}{8} \geq \frac{1}{2} - f,$$

which is always satisfied under the assumptions that $f > \underline{f}$. Hence, it is *a fortiori* satisfied for $p_2 < p^m$.

If B_1 rejected the offer, instead, B_2 would prefer buying from I rather than setting up a joint self-supply with the other buyer if: $CS(p_2) \geq CS(0) - f/2$, that is:

$$\frac{(1 - p_2)^2}{2} \geq \frac{1}{2} - f,$$

which amounts to:

$$p_2 \leq 1 - \sqrt{1 - f} \equiv p^*.$$

Note that p^* , which can be seen as a limit price, increases with f : the more expensive self-supply, the higher the price the incumbent can

charge without inducing the buyers to set up a joint-venture to produce the good themselves. In particular, $p^* \rightarrow (1 + c_I)/2 = p^m$ as $f \rightarrow \bar{f}$. It is easy to show that $p^* \geq c_I$ if (and only if)

$$f \geq 1 - (1 - c_I)^2 \equiv f_{UP}.$$

Consider now stage 1. B_1 observes the price p_2 and is able to anticipate the other buyer's choice. In particular, it knows that if $p_2 \leq p^*$, then B_2 will accept for sure the incumbent's offer, leaving B_1 obliged to accept whatever price $p_1 \leq p^m$ is offered. If, instead, $p_2 \in (p^*, p^m]$, then buyer B_1 anticipates that the second buyer will follow B_1 's decision and will accept the incumbent's offer only if the first buyer accepts. As a consequence, B_1 will accept the incumbent's offer if $p_1 \leq p^*$ (that is, if the price p_1 is sufficiently low to make it more convenient to buy from I than to establish joint production with B_2), and will reject it if otherwise.

We can now move to the incumbent's price-setting stage, and here we have to distinguish the two price regimes.

Uniform pricing If the incumbent cannot discriminate, it knows that the only way to avoid its clients 'integrating upstream' would be to set $p \leq p^*$. Clearly, profit maximisation will make it choose $p = p^*$. Note, however, that this is an equilibrium only if the incumbent makes positive profits, that is if $p^* \geq c_I$.

Therefore we have the following result:

Proposition 2.1 *Under uniform pricing, the equilibrium consists of:*

- if $f \in (f, f_{UP})$, firm I offers $p = c_I$, both buyers reject the offer and will self-supply;
- if $f \in [f_{UP}, \bar{f})$, firm I offers $p = p^*$, and both buyers accept the offer.

The threshold f_{UP} is below \bar{f} if (and only if) $c_I < c_1 \equiv 1 - \sqrt{4/7}$.

Price discrimination If the incumbent can price discriminate, it knows that it would be enough to offer p^* to one buyer to make sure that both buyers will accept its offers (since the other buyer would never find it convenient to self-supply, even if offered to buy at the monopoly price). Therefore the pair of prices $p_i = p^*$ and $p_j = p^m$ represents an equilibrium offer if (and only if) it allows the incumbent to make positive profits:

$$\pi_I(p^m, p^*) = \frac{(1 - c_I)^2}{4} + \left(1 - c_I - \sqrt{1 - f}\right) \sqrt{1 - f} \geq 0,$$

which after solving can be rewritten as:

$$f \geq 1 - \frac{(1 + \sqrt{2})^2 (1 - c_I)^2}{4} \equiv f_{PD}.$$

Note that $f_{PD} < f_{UP}$: under price discrimination the incumbent can offer p^* to one buyer only, while setting the monopoly price to the other buyer. Then, the incumbent can find it profitable to offer $p^* < c_I$ to one buyer, provided p^* is not too low and does not entail losses that are too large. This is the case if the fixed cost associated to self-supply is not too low. Therefore:

Proposition 2.2 *Under discriminatory pricing, the equilibrium consists of:*

- if $f \in (f, f_{PD})$, firm I offers $p = c_I$, both buyers reject the offer and will self-supply;
- if $f \in [f_{PD}, \bar{f})$, firm I offers $p_i = p^* = 1 - \sqrt{1 - f}$, $p_j = p^m = (1 + c_I)/2$, and both buyers accept the offer.

The threshold f_{PD} is below \underline{f} if $c_I < c_2 \equiv 1 - \frac{2}{\sqrt{2(1+\sqrt{2})^2 - 1}}$, with $c_2 > c_1$.

A similar logic can be applied to the case where there are N buyers (modifying the conditions on f appropriately). In fact, the decision of a buyer to accept the incumbent's offer makes it less beneficial for the other buyers to reject and form a joint-venture, by making the per-member cost larger. This externality is exploited by the incumbent who offers a decreasing price to a subset of buyers and the monopoly price to the remaining buyers. The lowest price p^* discourages the buyer from forming a full-size joint-venture. Hence, this buyer is better off accepting the monopolist's offer irrespective of the choice of the other buyers. It follows that the monopolist can elicit acceptance from another buyer by offering a higher price, as forming an organisation with $N - 1$ buyers involves higher per-member costs and is less beneficial. The same logic applies until, by bribing a sufficiently large group of buyers, the monopolist makes it unprofitable for the remaining ones to form a joint venture even if they are offered the monopoly price.

Welfare analysis To compare welfare when price discrimination is allowed and when, instead, the incumbent is restricted to use uniform prices, refer to Figure 2.1.

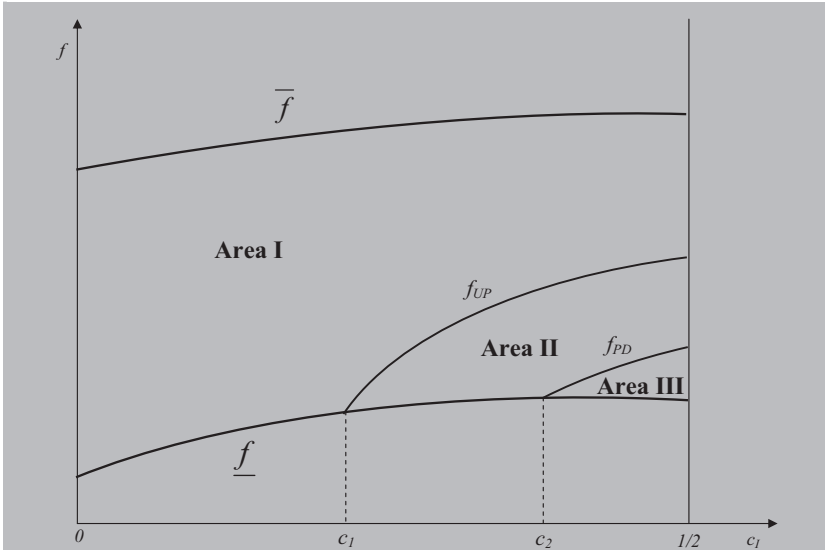


Figure 2.1. Welfare comparison.

For the parameter values corresponding to Area III, there will be self-supply under both price regimes, so welfare comparisons are immaterial.

For the parameter values corresponding to Area I, under uniform prices the incumbent offers a pair of prices (p^*, p^*) , whereas it offers (p^*, p^m) under discriminatory prices.

Therefore, under price discrimination the incumbent is able to extract higher surplus than under price uniformity. Note that this is welfare-detrimental, because it creates an allocative inefficiency relative to the case of uniform pricing: instead of buying at the lower price p^* , one buyer will buy at the monopoly price, which will entail a welfare reduction due to the deadweight loss. Note also that productive efficiency is not affected as both under price discrimination and under uniform pricing it is always the incumbent which produces.

In this case, then, price discrimination is welfare-detrimental not because it increases the scope for exclusion – in fact self-supply is discouraged irrespective of price discrimination – but because it allows to discourage self-supply in a way that is more profitable for the incumbent but more distortionary in terms of consumer surplus and total welfare.

Finally, let us consider Area II. In this region, there would be self-supply if firm *I* cannot price discriminate, and there would be 'exclusion' with prices (p^*, p^m) if it can. Welfare implications are *a priori* ambiguous here, since we do not know whether the allocative inefficiency caused by higher prices (p^*, p^m) is outweighed or not by possible productive efficiencies (recall that self-supply has a lower marginal cost but a positive fixed cost).

Let us analyse this case in detail. We want to see whether welfare under self-supply (which is what occurs under uniform pricing) is higher than welfare under price discrimination, that is:

$$W_{UP} = 2CS(0) - f \geq W_{PD} = CS(p^*) + CS(p^m) + \pi_I(p^m) + \pi_I(p^*).$$

Note that p^* was defined so that: $CS(p^*) = CS(0) - f/2$, implying that $2CS(0) = 2CS(p^*) + f$. The inequality $W_{UP} \geq W_{PD}$ can then be rewritten

as $CS(p^*) \geq CS(p^m) + \pi_I(p^m) + \pi_I(p^*)$, or:

$$\frac{(1-p^*)^2}{2} \geq \frac{3(1-c_I)^2}{8} + (p^* - c_I)(1-p^*),$$

which after substitution becomes:

$$12(1-f) - 8(1-c_I)\sqrt{1-f} - 3(1-c_I)^2 \geq 0,$$

which is solved for $f \leq 1 - \frac{(2+\sqrt{13})^2}{36}(1-c_I)^2 \equiv f_W$. Simple inspection of the expressions reveals that $f_W > f_{UP}$, which implies that in the region considered here $W_{UP} > W_{PD}$. Hence, in this region price discrimination is welfare-detrimental because it discourages self-supply.

The following proposition summarises the welfare analysis:

Proposition 2.3 *A ban on price discrimination is always welfare-beneficial (weakly so when self-supply takes place at the equilibrium irrespective of price discrimination).*

Discussion Strictly speaking, Innes and Sexton (1993)'s paper is not about exclusion of an efficient firm. In that paper self-supply would not increase productive efficiency because it entails the same marginal cost as the incumbent and, in addition, a fixed cost that the incumbent has already sunk. For this reason the price p^* that the incumbent must offer to discourage self-supply is above its marginal cost and it is profitable to offer such a price even under price uniformity. In other words, price

discrimination does not increase the scope of exclusion, as it happens in area I of Figure 2.1. However, this paper is probably the first to illustrate a mechanism which will be stressed by the subsequent literature on exclusionary conduct, namely that when entry entails scale economies, the incumbent can use a ‘divide-and-conquer’ strategy whereby a certain number of buyers are offered good terms of purchase, thereby reducing the scale available for entry (in this case, for establishing self-supply) and exerting a negative externality on the remaining buyers. This allows the incumbent to charge the remaining buyers with a much higher price. The notion that an incumbent firm can exclude more efficient rivals by exploiting this externality is the key idea of this book. A similar externality is central to the model of predation discussed in Chapter 1 (due to Fumagalli and Motta, 2013) and to the extensions of that model that rationalise conditional rebates (Section 2.4.3 of this chapter) and bundled rebates (Chapter 4). It is central to the paper by Karlinger and Motta (2012) where selective price cuts can exclude a more efficient rival, reviewed in the next section. It is also central to the literature on anti-competitive exclusive contracts, discussed in Chapter 3.

In fact, slight modifications in Innes and Sexton (1993) would render its results very similar to some of the results obtained by those papers. For instance, in the model discussed here mis-coordination doesn’t take place because price offers are publicly observed and buyers choose sequentially. But if buyers chose simultaneously or prices were secret, then there would be exclusionary equilibria where buyers end up paying the monopoly price simply by mis-coordination, as it happens in the paper by Karlinger and Motta (2012), and similarly to Segal and Whinston (2000a) (for the case where buyers simultaneously decide on exclusive dealing contracts).

Further, in Innes and Sexton’s model the ‘divide-and-conquer’ strategy is costly, but suppose we slightly reformulate the game analysed above as follows: (1a) the Incumbent makes offers to B_1 ; (1b) Buyer B_1 decides whether to accept or not; (2a) the Incumbent makes offers to B_2 ; (2b) Buyer B_2 decides whether to accept or not; (3) Buyers who have rejected I ’s offer decide on self-supply. In this case, we will obtain a result similar to the one obtained by Segal and Whinston (2000a) in the sequential version of their model, namely that exclusion will take place with both buyers paying the monopoly price. This is because B_1 anticipates that even if she rejects, the incumbent could always persuade B_2 to accept by offering her a price p^* , leaving B_1 with no incentive to

establish self-supply alone. Therefore, B_1 will be willing to accept the incumbent's offer even if she was offered p^m . (And B_2 will accept p^m because after the other buyer has accepted, alone it would not find it convenient to self-supply.)

Finally, Innes and Sexton's (1993) result reminds us of the predation model analysed in Chapter 1, with one buyer buying at a low price (to ensure that entry would not be profitable) and the other buying at a monopoly price.

2.4.2 (Explicit and Implicit) Selective Discounts*

In this section we will refer to the entry deterrence version of the model developed in Chapter 1, where the incumbent firm, I , has already sunk its entry costs, and a potential entrant, E , is considering entry. Firm E is more cost-efficient than firm I ($c_E = 0 < c_I < 1/2$), but it still needs to pay its entry cost, f . Moreover, we will focus on the case in which the incumbent and the entrant make *simultaneous* price offers to both buyers and buyers choose simultaneously.

Differently from the model of Chapter 1, we assume that buyers are asymmetric: instead of buying (at most) one unit each, there is a small buyer B^s which buys at most α units, and a large buyer B^l which buys at most $2 - \alpha$ units, with $\alpha \leq 1$. The asymmetry between the buyers is introduced to deal with the case of implicit discrimination, which would otherwise be meaningless. The case of symmetric buyers analysed in Chapter 1 (specifically, in Section 1.3.2.2) can be obtained as a special case by imposing $\alpha = 1$. Each buyer still has a unit valuation for the homogeneous product. To formalise the idea that selling to one buyer (at competitive prices) does not suffice for the entrant to be profitable, but two buyers do, we assume that:

$$\alpha c_I < (2 - \alpha)c_I < f < 2c_I. \quad (2.5)$$

In order to simplify the analysis and not to deal with corner solutions (see below) we also assume that $\alpha > f$. Therefore, we have:

$$f < \alpha \leq 1. \quad (2.6)$$

We shall consider three regimes, corresponding to different versions of the game. First, the case where firms cannot discriminate between buyers (*uniform pricing*). Second, the case of *explicit price discrimination*,

where firms can offer individualised prices to buyers whose characteristics are perfectly observable (for instance, they can offer a low price to the large buyer and a high price to the small buyer). Finally, the case where firms can only engage in *implicit price discrimination*, that is, where they are not able to observe whether a particular buyer is small or large, or they are not allowed to condition the price on such characteristics. In this case, price discrimination must satisfy a self-selection constraint (for instance, they will set a menu where the high price is offered in case of a small purchase and a low price for a large purchase, but buyers are free to choose the price/quantity combination they prefer).

The timing is as following:

1. Firms I, E make price offers simultaneously to buyers B_1 and B_2 . (Under uniform pricing, firms offer p_I and p_E . Under explicit discrimination, they can offer p_I^j, p_E^j , with $j = 1, 2$. Under implicit discrimination, they will offer price/quantity menus.)
2. Buyers simultaneously decide from whom to buy (and are committed to their choice).
3. Firm E decides whether to enter (pay f) or not.
4. Transactions are made.
5. (If E got orders from a buyer B_i at stage 2, but later does not enter, then stages 1–4 are repeated for B_i .)

2.4.2.1 The base model, with simultaneous offers and price uniformity*

We first consider the case where firms cannot price discriminate across buyers; that is, they are constrained to set $p_k^1 = p_k^2$. Similarly to the case of simultaneous offers discussed in Chapter 1 (see in particular Section 1.3.2.2), there are two equilibria which always arise. Exclusion arises because of scale economies: the entrant needs to make a certain amount of profits to be able to recover its entry cost. Serving only one buyer would not be sufficient. This introduces an externality among buyers. If one buys from the incumbent, the other will have no incentive to choose the entrant, as it knows that its demand alone is not large enough to induce entry. This creates the scope for coordination failures and explains why there is a Nash equilibrium where both buyers buy from I , although they may be offered prices as high as the monopoly price by the incumbent and lower prices by the entrant: a unilateral

deviation would not leave a buyer better off – as she would not be able to trigger entry, and would end up buying from the incumbent later anyway.

Proposition 2.4 (*Uniform pricing, asymmetric buyers*) *The game admits two types of equilibria (they both exist for all parameter values).*

- *Exclusionary (mis-coordination) equilibria.* Firm I sets a price $p_I \in [c_I, f/(2 - \alpha)]$, firm E sets $p_E \leq p_I$, all buyers buy from I and entry will not occur.
- *Entry equilibria.* Firm E sets $p_E^* \in [f/2, c_I]$, firm I sets $p_I^* \in [p_E^*, f]$, both buyers buy from E and entry will occur.

Proof. The proof is by backward induction. At the buyer stage: If $p_I < p_E$, there is a unique equilibrium where both buyers choose the incumbent firm. If $p_E < p_I < f/(2 - \alpha)$ there are two equilibria in the buyers' game: one, where both buyers choose firm E and the other, where both buyers choose firm I for the typical mis-coordination argument. Given that buyer B^S buys from I , buyer B^J has no incentive to deviate and choose E . If she did so, entry would not follow as her demand alone is insufficient to make entry profitable: $\pi_E = (2 - \alpha)p_E < f$ since $p_E < f/(2 - \alpha)$. *A fortiori*, the small buyer will have no incentive to deviate given that the large buyer buys from I . If instead $p_I > f/(2 - \alpha)$ and $p_E < p_I$ but sufficiently close to it, then there cannot be a continuation equilibrium where both buyers choose firm I : the large buyer would deviate and trigger entry.

We can now characterise the equilibrium price configurations. Consider first the **mis-coordination equilibria**. The pair of prices $p_I^* = f/(2 - \alpha)$ and $p_E^* \leq p_I^*$ is sustained by the continuation equilibrium where both buyers choose the incumbent following any bid where $p_E \leq p_I = f/(2 - \alpha)$: firm E has no incentive to increase its price as both buyers will choose the incumbent; firm I has neither an incentive to decrease its price (it would lower profits), nor an incentive to increase its price as it would lose at least one buyer. More generally, the prices $p_I^* = p < f/(2 - \alpha)$ and $p_E^* \leq p_I^*$ are sustained by having both buyers choosing the incumbent following any bid where $p_E \leq p_I = p$, while both buyers choosing the entrant following any bid $p_I > p$ and $p_E \leq p_I$. These continuation equilibria ensure that the incumbent has no incentive to deviate and bid a price above p because it would lose all

buyers; also, the entrant has no incentive to change its bid because this would not allow it to attract buyers nor to make entry profitable. Finally, a mis-coordination equilibrium where $p_I > f/(2 - \alpha)$ does not exist. Firm E would have an incentive to deviate and slightly undercut the incumbent: by getting the large buyer it will have sufficient profits to cover fixed costs.

Next, consider **entry equilibria**. A price $p_E > c_I$ cannot be an equilibrium as the incumbent would undercut and obtain all buyers. Firm E cannot enter the market if it bids a price $p_E < f/2$ either: the demand of both buyers is not enough to cover the entry costs. Equilibria where $p_E^* = p \in [f/2, c_I]$ and $p_I^* = p$ are sustained by having both buyers choose the entrant following any bid where $p_E < p_I$. The entrant cannot deviate by increasing its price as it would lose all orders. In turn, the incumbent is indifferent between p and any higher price because no buyer would patronise it in any case; instead, it captures both buyers by decreasing its price but it would not break even as the deviation price would be below c_I . Finally, there are also entry equilibria where $p_I > p_E^* = p \in [f/2, c_I]$. They are sustained by having both buyers choose the entrant following any bid where $p_I > p_E = p$ and both buyers choosing the incumbent following any bid where $p < p_E \leq p_I$. The latter ensures that firm E cannot increase its payoff by increasing the price and setting it equal or lower than the incumbent's because it would lose all the buyers. ■

2.4.2.2 Explicit price discrimination*

Assume now that both firms can price discriminate. (Since we have only two buyers, this could be interpreted as both first-degree and third-degree discrimination.) The following Proposition shows that the exclusionary (mis-coordination) equilibria always exist for all parameter values, whereas the entry equilibria exist only for a subset of the values. This is because the incumbent can exploit a 'divide-and-conquer' strategy: since the entrant needs both buyers, the incumbent can make an aggressive below-cost price offer to one buyer while recovering profits on the other, and *vice versa*: the entrant needs to set a price which is low enough on both buyers to prevent the incumbent from deviating by using such a strategy. This is profitable only if the efficiency gap is large enough.

Proposition 2.5 (*Explicit discrimination, asymmetric buyers*) The game admits two types of equilibria.

- *Exclusionary (mis-coordination) equilibria.* Firm *I* sets a price pair (p_I^s, p_I^l) with the highest prices being $p_I^s = f/\alpha$, $p_I^l = f/(2 - \alpha)$, firm *E* sets $p_E^j \leq p_I^j$, ($j = s, l$), all buyers buy from *I* and entry will not occur.
- *Entry equilibria.* If $f \leq f^{(expl)} \equiv 4c_I/3$ then the highest pair of prices that firm *E* offers are: $p_E^{s(expl)} = (2c_I - f)/\alpha < c_I$, $p_E^{l(expl)} = (2c_I - f)/(2 - \alpha) < c_I$, firm *I* sets $p_I^j = p_E^j$ with $j = s, l$, both buyers buy from *E* and entry will occur. If $f > f^{(expl)}$, no entry equilibrium exists.

Proof. As for **exclusionary equilibria**, even under price discrimination they arise because of the usual mechanism of mis-coordination. Consider for instance a candidate equilibrium where $p_E^s \leq p_I^s \leq f/\alpha$, $p_E^l \leq p_I^l \leq f/(2 - \alpha)$ and both buyers buy from *I*. Given that buyer B_j buys from *I* at price p_I^j , there would be no incentive for buyer B_i to deviate and buy from *E* even if p_E^i is lower than p_I^i , as the deviation would not trigger entry. Under the appropriate continuation equilibria, sellers would not have an incentive to deviate either. If firm *E* further decreases its prices, it would not switch to a continuation equilibrium where buyers select the entrant. And provided that by increasing its prices there would be a switch to a continuation equilibrium where buyers buy from the entrant, firm *I* would not have an incentive to increase prices either. (The equilibrium price configurations need to satisfy the condition that $\pi_I(p_I^s, p_I^l) \geq 0$ so as to ensure that the incumbent makes positive profits and does not want to deviate. That condition is always satisfied by $p_I^s = f/\alpha$ and $p_I^l = f/2 - \alpha$ as both prices are greater than c_I by assumption.)

Let us now turn to **entry equilibria**. For an entry equilibrium to exist, we need to find a pair (p_E^s, p_E^l) which is immune to deviations by the incumbent. Since the incumbent could block entry *either* (i) by selling to the small buyer at a price $p_I^s < p_E^s$ while serving the large buyer at the highest possible price which does not trigger entry, $\bar{p}_I^l = f/(2 - \alpha)$; or (ii) by selling to the large buyer at a price $p_I^l < p_E^l$ while recouping losses by serving the small buyer at the highest possible price which does not trigger entry $\bar{p}_I^s = f/\alpha$, at an entry equilibrium *both* of the entrant's prices must be such that no profitable undercutting by the incumbent could take place. Formally, under the possible deviations (i) and (ii) the

incumbent's offers must satisfy $\pi_I \geq 0$:

$$(i) \quad \alpha p_I^s + (2 - \alpha) \frac{f}{2 - \alpha} \geq 2c_I,$$

$$(ii) \quad (2 - \alpha)p_I^l + \alpha \frac{f}{\alpha} \geq 2c_I,$$

The incumbent's most aggressive price offers would therefore be respectively:

$$\underline{p}_I^{s(\text{expl})} = \frac{2c_I - f}{\alpha} < c_I; \quad \underline{p}_I^{l(\text{expl})} = \frac{2c_I - f}{2 - \alpha} < c_I,$$

For an entry equilibrium to exist, the entrant must set prices which are immune to both deviations, namely $p_E^s = \underline{p}_I^s$, and $p_E^l = \underline{p}_I^l$. Therefore, an entry equilibrium will exist if and only if $\pi_E(\underline{p}_I^s, \underline{p}_I^l) = \alpha \underline{p}_I^s + (2 - \alpha) \underline{p}_I^l \geq f$. By substitution, an entry equilibrium exists if and only if $f \leq 4c_I/3 \equiv f^{(\text{expl})}$. ■

Note that under explicit discriminatory pricing, the highest prices and profits that the incumbent can make at the *exclusionary equilibrium* are (weakly) higher than under uniform pricing. In particular, the incumbent can identify the small buyer and offer her the higher price f/α while offering the large buyer the lower price $f/(2 - \alpha)$. Any price to the large buyer higher than $f/(2 - \alpha)$ would not be an equilibrium price: the large buyer would have an incentive to address the entrant, as now this would trigger entry. This is why at equilibrium the large buyer benefits from a lower price than the small buyer.

Further, when the entry equilibrium exists, it is characterised by the small buyer paying a higher price than the large buyer.

Finally, in the case of sequential buyers (see Chapter 1) entry equilibria exist under a larger parameter space than in the case of simultaneous discriminatory offers ($f < 3c_I/2$ instead of $f < f^{(\text{expl})} \equiv 4c_I/3$). This is because in the former case, in order to avoid exclusion the entrant needs to price quite aggressively on the first buyer, while it will manage to charge the price c_I to the second buyer. Instead, when offers are simultaneous like in this chapter, the entrant needs to offer prices below c_I to *both* buyers, so as to discourage the incumbent from engaging in divide-and-conquer deviations.

2.4.2.3 Implicit (second-degree) price discrimination*

In this section, we shall consider the case where firms are not able to target specific buyers, but have to rely on implicit (or second-degree) price

discrimination. Each firm can offer a menu of the type $\{(p_i^s, q_i^s), (p_i^l, q_i^l)\}$ under which the buyer can choose whether to purchase a certain number of units q_i^s at the price p_i^s or certain number of units q_i^l at the price p_i^l .

Proposition 2.6 (*Implicit discrimination, asymmetric buyers*) *The game admits two types of equilibria.*

- *Exclusionary (mis-coordination) equilibria.* Firm I sets a menu $\{(p_I^s, \alpha), (p_I^l, 2 - \alpha)\}$ with the highest prices being $p_I^s = f/\alpha$ and $p_I^l = f/(2 - \alpha)$, firm E sets $\{(p_E^s, \alpha), (p_E^l, 2 - \alpha)\}$ with $p_E^j \leq p_I^j$, ($j = s, l$), all buyers buy from I and entry will not occur.
- *Entry equilibria.* If $f \leq f^{(impl)}$ then firm E sets a menu $\{(p_E^s, \alpha), (p_E^l, 2 - \alpha)\}$, with the highest pair of prices being $p_E^{s(impl)} < c_I$ and $p_E^{l(impl)} < c_I$, firm I sets the menu $\{(p_I^s = p_E^s, \alpha), (p_I^l = p_E^l, 2 - \alpha)\}$, both buyers buy from E and entry will occur. If $f > f^{(impl)}$, no entry equilibrium exists. It should be noted that $p_E^{j(impl)} \geq p_E^{j(expl)}$ with $j = s, l$.

Proof. For the *exclusionary equilibria*, the proof is identical as the previous cases, relying on mis-coordination between buyers. Additionally, the self-selection constraints must be satisfied: both the large and the small buyer must prefer to purchase the price and quantity designed for them, rather than mimicking the other buyer. It turns out that the incumbent can enjoy the same maximum prices as under explicit discrimination: the large buyer clearly would not have an incentive to buy fewer units at a higher price; and the small buyer would not increase its surplus if instead of buying at the higher price f/α she mimicked a large buyer and bought $2 - \alpha$ units at the lower price: $\alpha(1 - f/\alpha) = \alpha - (2 - \alpha)[f/(2 - \alpha)]$.

To find *entry equilibria*, we follow the same logic as under explicit discrimination, and we identify the incumbent's most aggressive price offers, with the additional requirement that price offers satisfy a self-selection constraint. Formally, then, the incumbent's offers must satisfy:

$$(i') \quad \alpha p_I^s + (2 - \alpha)\bar{p}_I^l = 2c_I, \text{ s.t.: } (2 - \alpha)(1 - \bar{p}_I^l) \geq \alpha(1 - \underline{p}_I^s);$$

$$\bar{p}_I^l \leq f/(2 - \alpha); \underline{p}_I^s \geq 0.$$

$$(ii') \quad (2 - \alpha)\underline{p}_I^l + \alpha\bar{p}_I^s = 2c_I, \text{ s.t.: } \alpha(1 - \bar{p}_I^s) \geq \alpha - (2 - \alpha)\underline{p}_I^l;$$

$$\bar{p}_I^s \leq f/\alpha; \underline{p}_I^l \geq 0.$$

Consider first (i'). Imagine the incumbent makes the same most aggressive price offer as under explicit discrimination while recouping on the other buyer: $\underline{p}_I^s = (2c_I - f)/\alpha$ and $\bar{p}_I^l = f/(2 - \alpha)$. By replacing these values in the incentive constraint of the large buyer, we see that it is satisfied if and only if $f \leq 1 + c_I - \alpha$. For higher values of f , we have to impose that self-selection is satisfied. From $(2 - \alpha)(1 - \bar{p}_I^l) \geq \alpha(1 - \underline{p}_I^s)$ we obtain that, given \underline{p}_I^s the highest price that the incumbent can charge to the large buyer still satisfying the self-selection constraint is:

$$\bar{p}_I^l = \frac{2(1 - \alpha) + \alpha \underline{p}_I^s}{2 - \alpha}.$$

By substituting into (i') and solving we obtain the lowest price the incumbent could offer to the small buyer:

$$\underline{p}_I^{s(impl)} = \frac{c_I - 1 + \alpha}{\alpha} > \underline{p}_I^{s(expl)} = \frac{2c_I - f}{\alpha}.$$

Not surprisingly, when the incumbent cannot make targeted offers to buyers it cannot be as aggressive as in the case of explicit price discrimination.

Consider now condition (ii'). One can check that the prices which apply under explicit discrimination ($\underline{p}_I^l = \frac{2c_I - f}{2 - \alpha}$, $\bar{p}_I^s = \frac{f}{\alpha}$) never satisfy the self-selection constraint of the small buyers. In order to find the pair of prices which would be compatible with self-selection by the small buyer, set $\alpha(1 - \bar{p}_I^s) = \alpha - (2 - \alpha)\underline{p}_I^l$, whence:

$$\bar{p}_I^s = \frac{2(1 - \alpha)}{\alpha} \underline{p}_I^l.$$

By substituting into (ii'') and solving, one obtains that the most aggressive price offer the incumbent can make to the large buyer, while keeping the self-selection constraint of the small buyer satisfied, is:

$$\underline{p}_I^{l(impl)} = \frac{c_I}{2 - \alpha} > \underline{p}_I^{l(expl)} = \frac{2c_I - f}{2 - \alpha}.$$

Again, the incumbent cannot be as aggressive as in the case of explicit price discrimination.⁴⁴

⁴⁴ Note that off the equilibrium path the incumbent may use a price schedule whereby a lower price is offered for a smaller number of units and a higher price for a larger number

To sum up, we have to distinguish two cases. Case I: $f \leq 1 - \alpha + c_I$. In this case, to be immune to both types of incumbent's deviations, the entrant's prices cannot exceed:

$$p_E^s = \underline{p}_I^{s(expl)} = \frac{2c_I - f}{\alpha}; \quad p_E^l = \underline{p}_I^{l(impl)} = \frac{c_I}{2 - \alpha}.$$

It can be easily verified that the candidate equilibrium offers by the entrant are themselves incentive-compatible and both buyers purchase the offer intended for them.

Finally, we have to check under which conditions the entrant can profitably make those offers. The entrant breaks even under the candidate equilibrium prices if:

$$\pi_E(\underline{p}_I^{s(expl)}, \underline{p}_I^{l(impl)}) = \alpha \frac{2c_I - f}{\alpha} + (2 - \alpha) \frac{c_I}{2 - \alpha} - f \geq 0.$$

This is satisfied for $f \leq \frac{3c_I}{2} \equiv f_1^{(impl)}$. Note that if $\alpha > 1 - c_I/2$, then $f_1^{(impl)} > 1 - \alpha + c_I$. Hence, entry equilibria always exist under the condition that characterises Case I. Instead, if $\alpha \leq 1 - c_I/2$, then in Case I entry equilibria exist if $f \leq f_1^{(impl)}$.

Let us turn to Case II: $f > 1 - \alpha + c_I$. In this case the entrant's price cannot exceed:

$$p_E^s = \underline{p}_I^{s(impl)} = \frac{c_I - 1 + \alpha}{\alpha}; \quad p_E^l = \underline{p}_I^{l(impl)} = \frac{c_I}{2 - \alpha}.$$

It is easy to verify that also these offers are incentive-compatible. Turning to profitability, entry equilibria exist if:

$$\pi_E(\underline{p}_I^{s(impl)}, \underline{p}_I^{l(impl)}) = \alpha \frac{c_I - 1 + \alpha}{\alpha} + (2 - \alpha) \frac{c_I}{2 - \alpha} - f \geq 0.$$

This is satisfied for $f \leq 2c_I - 1 + \alpha \equiv f_2^{(impl)}$. Note that if $\alpha \leq 1 - c_I/2$, then $f_2^{(impl)} \leq 1 - \alpha + c_I$. Hence, entry equilibria never exist under the condition that characterises Case II. Instead, if $\alpha > 1 - c_I/2$, then in Case II entry equilibria exist as long as $f \leq f_2^{(impl)}$.

Putting the two cases together, we can conclude that entry equilibria exist if $f \leq f^{(impl)}$ where $f^{(impl)} \equiv 3c_I/2$ if $\alpha \leq 1 - c_I/2$ and $f^{(impl)} \equiv 2c_I - 1 + \alpha$ if $\alpha > 1 - c_I/2$. ■

of units, in order to attract the small buyer. This is the opposite of a quantity discount. However, along the equilibrium path a quantity discount scheme is offered whereby the unit price paid for a small order is higher than for a large order.

Note that under explicit discrimination the prices the entrant has to make to ensure immunity from deviations are (weakly) higher than those it has to make under explicit discrimination. Hence, profitability will be (weakly) higher and entry equilibria will be more likely to exist. The following Proposition expresses this formally.

Proposition 2.7 (*Comparison among entry equilibria*) *Entry equilibria exist for any feasible parameter values under uniform pricing. Under price discrimination there exist feasible parameter values for which entry equilibria do not exist any longer. The set of parameters that sustain entry equilibria is larger when price discrimination is implicit than explicit.*

Proof. It follows in a straightforward way from the comparison between $f^{(impl)}$ and $f^{(expl)}$. When $\alpha \leq 1 - c_I/2$, $f^{(impl)} = 3c_I/2 > 4c_I/3$. When $\alpha > 1 - c_I/2$, $f^{(impl)} = 2c_I - 1 + \alpha > 4c_I/3$. ■

Note an important difference between the model with simultaneous offers to all the buyers that we are proposing in this section and the model with sequential offers developed in Chapter 1 and that we will study again in the next sections dealing with conditional discounts. When offers are sequential, below-cost prices arise at the equilibrium. When instead offers are simultaneous, below-cost prices are part of a divide-and-conquer strategy that allows the incumbent to break entry equilibria that would otherwise arise under uniform prices, thereby reducing the set of achievable entry equilibria. However, equilibrium prices are not necessarily below cost in the model with simultaneous offers. Indeed, exclusion relies on buyers' coordination failures and it is compatible with prices as high as the monopoly price.

Equilibrium selection

Unlike the model with sequential buyers, where the equilibrium was unique for any given combination of parameter values, the model with simultaneous buyers admits multiple equilibria. Indeed, we know, for instance, that under uniform pricing for all admissible sets of values, both the exclusionary and the entry equilibria are possible. In this section, we study whether there exist refinements which select among these equilibria. In what follows, we consider two standard criteria for equilibrium selection, namely coalition-proofness and risk-dominance. Under the former, it turns out that only entry equilibria would survive; under the latter, instead, whether exclusionary or entry equilibria are selected depends on the equilibrium prices prevailing in the exclusionary equilibria.

Coalition-proof Nash equilibria The coalition-proof Nash equilibrium concept is due to Bernheim et al. (1987). In a coalition-proof equilibrium there is no coalition of agents for which a self-enforcing deviation would make all of its members better off. In other words, suppose that different groups of buyers can coordinate their purchases. Then a coalition-proof equilibrium exists if there is no such group of buyers which would improve its payoff by (jointly) deviating. In what follows, we show that the joint deviation of both buyers together would generate critical mass for the entrant and would disrupt any exclusionary equilibrium. Instead, all entry equilibria would satisfy the coalition-proofness requirement. Therefore, if we applied this selection criterion only the entry equilibria would exist.

Proposition 2.8 (*Coalition-proofness, uniform pricing*) *Under uniform pricing, both exclusionary and entry equilibria always exist (see Proposition 2.4). (a) The exclusionary equilibria are not coalition-proof. (b) The entry equilibria are coalition-proof.*

Proof. (a) To sustain a mis-coordination equilibrium, the continuation equilibria must be such that buyers mis-coordinate on I as long as $p_E \leq p_I$. Consider the case where $p_E \in [c_I, p_I)$, and suppose the coalition of our two buyers deviates to choose the entrant instead of the incumbent. Since $2c_I > f$, firm E would be profitable, thus improving the payoffs of both members of the coalition, who would pay a lower price. After switching to E , given that $p_E < p_I$, neither individually nor jointly could buyers improve their payoffs by switching back to I . This implies we have found a coalition for which it is profitable to deviate from the continuation equilibria identified above. In turn, this means that at the price setting stage of the game, firm E will always want to slightly undercut firm I ; since the lowest price firm I can offer is c_I , and since $2c_I > f$, imposing coalition-proofness at the supplier choice stage of the game will result in a profitable deviation by E at the price stage of the game. No exclusionary equilibria can survive firm E 's deviation.

(b) By construction of the entry equilibria, they are immune against all possible price offers by the incumbent aimed at inducing a deviation by one or both buyers in order to deprive the entrant of the minimum profits it needs. This implies there exist no self-enforcing deviations by any coalition of buyers in the continuation subgames. Both buyers

buying from E at $p_E = p_I$ is indeed a Nash equilibrium at the price stage of the game. ■

Although we prove it here for the uniform pricing regime, the same result applies to the other cases: entry equilibria are coalition-proof, mis-coordination equilibria are not.

Risk-dominance Another commonly used criterion of equilibrium selection is risk-dominance. The application in this game is not completely straightforward because it involves both players in the first stage (sellers) and in the following stage (buyers). Let us focus on the buyers' stage, and see if there are criteria which may select a particular equilibrium. For simplicity, let us also assume that buyers are symmetric ($\alpha = 1$). Consider for instance a game where buyers are offered $p_E \in [f/2, c_I]$ and $p_I \in [c_I, f]$, as in Table 2.1. Note that if a buyer chooses E but the other chooses I , entry will not occur and the buyer will have to buy from the incumbent at a later stage, and pay f (recall that firm I cannot charge a higher price without inviting entry).

Table 2.1. *Buyers' payoffs*

B_1, B_2	(order from) E	(order from) I
(order from) E	$1 - p_E, 1 - p_E$	$1 - f, 1 - p_I$
(order from) I	$1 - p_I, 1 - f$	$1 - p_I, 1 - p_I$

The game admits the exclusionary equilibrium (I, I) where both buyers buy from I and the entry equilibrium (E, E) where both buyers buy from E . Clearly, *Pareto-dominance* would select the entry equilibrium. However, risk-dominance would not unambiguously select a particular equilibrium. To find the risk-dominant equilibrium, define the variables L_{Ki} , with $K = E, I$ and $i = 1, 2$. For instance, L_{E1} is the loss avoided by player B_1 when rightly playing E rather than wrongly playing I when B_2 chooses the equilibrium play E .

$$L_{E1} = (1 - p_E) - (1 - p_I) = p_I - p_E$$

$$L_{E2} = (1 - p_E) - (1 - p_I) = p_I - p_E$$

$$L_{I1} = (1 - p_I) - (1 - f) = f - p_I$$

$$L_{I2} = (1 - p_I) - (1 - f) = f - p_I$$

The equilibrium (E, E) risk-dominates (I, I) if $(L_{E1})(L_{E2}) > (L_{I1})(L_{I2})$, that is, if $(p_I - p_E)^2 > (f - p_I)^2$, which can also be rewritten as $p_I > (f + p_E)/2$.

In other words, the entry equilibrium is selected by the risk-dominance criterion only if the price corresponding to the particular exclusionary equilibrium assumed is sufficiently high, and more specifically above the mid-point between the entry price p_E and the highest price to be paid when making a ‘mistake’ and being the only one buying from the entrant, f .⁴⁵

2.4.3 Conditional Rebates to Target the Contestable Demand*

In this section, as in the base model of Chapter 1, we consider two identical buyers B_1 and B_2 . Differently from that model though, each buyer demands *two* units of a product. More precisely, each buyer demands one unit from the incumbent *for sure* (for example, because of switching costs, lock-in effects, or because for a large proportion of final consumers this is a must-buy brand), whereas the second one is *contestable*, i.e. it can be bought either from the incumbent or from the rival seller. Buyers have valuation equal to one for each unit of the product. The incumbent and the rival sell homogeneous products. Competition for the two suppliers takes place sequentially, with firms making price offers to B_1 first and then, after this buyer has decided from whom to buy, firms make price offers to B_2 . We will refer to the entry deterrence interpretation of the base model of Chapter 1, thereby assuming that the incumbent has constant marginal cost $c_I > 0$ while the rival has zero marginal cost, but it has not sunk the entry cost f yet when competition for the first buyer starts, with $c_E = 0 < c_I < 1/2$ and:

$$c_I < f < 2c_I \quad (2.7)$$

As we have highlighted repeatedly, such a cost structure refers also to the case in which the rival is already in the market but incurs a cost $f > c_I$ to produce the first contestable unit and a cost equal to zero to produce the second contestable unit.

We model conditional rebates in a very simple way. The incumbent offers each buyer $i = 1, 2$ a price schedule (p_{Ii}, R_i) where p_{Ii} is the unit

⁴⁵ For instance, if we assumed $p_E = c_I$ (which seems a natural benchmark) then $p_I > (f + c_I)/2$.

price and R_i is the discount given if the buyer buys both units from the incumbent. The rival offers the price p_{Ei} to buyer $i = 1, 2$: it does not need to offer conditional discounts as it only competes for the contestable unit of each buyer.

It follows that buyer B_i pays the total price $2p_{Ii} - R_i$ if she buys both units from the incumbent, and the total price $p_{Ii} + p_{Ei}$ if she buys the contestable unit from the entrant. She will buy from the entrant if (and only if) $p_{Ei} \leq p_{Ii} - R_i$, where $p_{Ii} - R_i$ can be denoted as the incumbent's *effective price* for the contestable unit. The effective price is the price that the rival has to match so as to attract the contestable unit.

The price schedule offered by the incumbent can be interpreted in different ways: for instance, as a quantity discount whereby the buyer needs to buy (strictly) more than one unit to qualify for the discount; or as a market-share discount, whereby the buyer needs to purchase (strictly) more than 50 per cent of her requirement from the incumbent to qualify for the discount; or as an exclusivity discount, whereby the buyer needs to purchase all of her requirement from the incumbent to qualify. Whatever the form, the point is that in this environment conditionality allows the incumbent to discriminate the price across different units and to target the discount to the contestable part of buyers' demand.

Moreover, note that in order to properly compute the effective price, one should allocate to the contestable units the *entire* discount that the buyer loses when she buys from the rival and does not qualify for the discount. If the discount is expressed as a lump-sum discount as in this case, the total discount lost is R_i and the effective price is $p_{Ii} - R_i$. An equivalent discount scheme would be the one in which a per-unit discount (or a percentage discount) is applied to the incremental units exceeding the threshold. In this simple example with two total units, the incremental per-unit discount would amount to R_i (or to $R_i/p_{Ii}\%$) if the buyer buys more than one unit from the incumbent, or more than 50 per cent of her requirement. Another equivalent discount scheme might be retroactive so that, once the buyer qualifies for the discount, the discount is applied to all the units that the buyer purchases from the incumbent. In this case, a retroactive discount equivalent to the previous ones would be such that the unit discount $R_i/2$ – or $R_i/(2p_{Ii})\%$ – applies to both units if the buyer purchases more than one unit from the incumbent, or more than 50 per cent of her requirement. Under such a retroactive rebate scheme, the total discount that the buyer would lose by addressing

the rival is indeed $2(R_i/2) = R_i$ and this is the discount that should be allocated to the contestable units to correctly identify the effective price.

The game is as follows:

1. First period.

- (a) Firms I, E simultaneously offer (p_{I1}, R_1) and p_{E1} to buyer 1.
- (b) Buyer 1 decides from whom to buy and commits to her choice.
- (c) Firm E decides whether to enter (and pay f) or not.
- (d) Transactions take place. If E got the order from buyer 1 but did not enter, buyer 1 purchases from I at the offered price (p_{I1}, R_1) .⁴⁶

2. Second period.

- (a) Firms simultaneously offer (p_{I2}, R_2) and p_{E2} to buyer 2.
- (b) Buyer 2 decides from whom to buy and commits to her choice.
- (c) If it has not entered yet, firm E decides whether to enter (and pay f) or not.⁴⁷
- (d) Transactions take place. If E got the order from buyer 2 but did not enter, buyer 2 purchases from I at the offered price (p_{I2}, R_2) .

Proposition 2.9 *Equilibria of the game are as follows:*

- **Exclusion** *If $f > 3c_I/2$, then the first buyer buys both units from the incumbent at a price $p_{I1}^* = 1$ and receives a discount $R_1^* = 1 - f + c_I$; the second buyer buys both units from the incumbent at the price $p_{I2}^* = 1$ and receives a discount $R_2^* = 1 - f$; the entrant does not enter the market.*
- **Entry/Expansion** *If $f \leq 3c_I/2$, then the entrant enters the market. Both buyers buy the non-contestable unit from the incumbent at the price $p_{I1}^* = p_{I2}^* = 1$ and the contestable unit from the entrant. The first buyer pays the price $p_{E1}^* = 2c_I - f$, while the second buyer pays $p_{E2}^* = c_I$.*

⁴⁶ The results would not change if we assumed that the buyer whose order remains unfulfilled is forced to buy from the incumbent which would then charge the monopoly price.

⁴⁷ Allowing the entrant to enter also at the end of the second period only affects the 'maximum' effective price that firm I could charge to the second buyer after it has already served the first buyer. Allowing for a second chance of entry implies that instead of charging the effective price $p = 1 > f$, the incumbent will charge the 'limit price' $p = f$ (if the price was higher, the entrant would undercut the incumbent and enter). Note that

Proof. The incumbent can govern the effective price for the contestable units through the choice of the discount R_i . Hence, it is optimal for the incumbent to set $p_{I1}^* = p_{I2}^* = 1$ and extract the entire surplus that buyers derive from the non-contestable units. The effective price for the contestable unit of buyer i is, then, $1 - R_i$.

We now proceed by backward induction. Let us analyse the second period, when suppliers make price offers to B_2 . If firm E did not enter at the end of the first period, then it would still have the chance to enter by paying the fixed cost f . Therefore, its cost to supply the contestable unit to B_2 is f . Since the marginal cost of the incumbent is $c_I < f$, at the equilibrium the incumbent secures the contestable unit of B_2 by setting the effective price $1 - R_2^* = f$. The discount offered to B_2 is, then, $R_2^* = 1 - f$. If firm E did enter at the end of period one, then its cost to supply the contestable unit to B_2 equals $c_E = 0$, which is lower than the incumbent's marginal cost, c_I . At equilibrium, the entrant wins B_2 's contestable unit by setting $p_{E2}^* = c_I$.

At the end of the first period, firm E takes the entry decision. If B_1 chose the incumbent, then firm E anticipates that the revenues it can collect from B_2 are insufficient to cover the entry cost: $c_I < f$. Then it decides not to enter the market.

If B_1 chose to buy from firm E , then entry is profitable if the revenues collected in the first period together with the ones that firm E anticipates to obtain in the second period are large enough to cover the entry cost: $p_{E1} + c_I \geq f$.

At the beginning of the first period, the two suppliers compete for the contestable unit of B_1 . The minimum price at which firm E is willing to supply B_1 is $\tilde{p}_{E1} = f - c_I$. The minimum effective price that the incumbent is willing to offer for B_1 's contestable unit is $1 - \tilde{R}_1 = 2c_I - f$: if the incumbent supplies B_1 's contestable unit, then entry will not follow and the incumbent's total profits amount to $2 - R_1 - 2c_I + 2 - 1 + f - 2c_I$; if the entrant supplies B_1 's contestable unit, the incumbent sells only the non-contestable units making total profits equal to $2 - 2c_I$.

The incumbent manages to win the contestable unit of the first buyer if (and only if) $2c_I - f < f - c_I$, that is, if (and only if) $f > 3c_I/2$. When this condition is satisfied, at the equilibrium the incumbent's effective

by assuming that entry is possible also in the second period we make it more difficult for exclusion to take place, since after having deterred entry in the first period, the incumbent is not able to set the monopoly price 1 but only the lower price f .

price for the first buyer is $1 - R_1^* = f - c_I$ which entails a discount $R_1^* = 1 - f + c_I$.

If instead $f \leq 3c_I/2$, then it is the entrant that wins the contestable unit of the first buyer by offering the equilibrium price $p_{E1}^* = 2c_I - f$. ■

Lemma 2.10 *At the exclusionary equilibrium, the incumbent charges an effective price for the contestable unit of the first buyer which is below its marginal cost: $1 - R_1^* = f - c_I < c_I$. However, it does not make overall losses on the first buyer, since the profit earned on the non-contestable unit outweighs the loss made on the contestable one.*

Proof. Below-cost pricing on the contestable unit follows from the assumption that $f < 2c_I$. Total profits made on the first buyer amount to $2 - R_1^* - 2c_I = 1 + f - 3c_I$, which is positive if (and only if) $f > 3c_I - 1$. This condition is always satisfied under our assumption that $c_I < 1/2$, which implies that $3c_I - 1 < c_I < f$. ■

2.4.4 Discounts Conditional on Exclusivity*

As above, we assume in this section that there exist two identical buyers, B_1 and B_2 , but we assume now that each buyer's demand is elastic and is given by $D(p) = 1 - p$. This assumption is important because it allows us to compare the results obtained when suppliers offer two-part tariffs and when they offer linear prices, and to highlight the role of the loyalty requirement in the contract offered to buyers. The buyers' demand is entirely contestable.

We start the analysis considering the case in which, in the first period, the incumbent and the entrant simultaneously offer B_1 a pricing scheme $t_j^1(q_1) = c_j q_1 + T_j^1$ with $j = I, E$ in which the linear component of the tariff is equal to the supplier's marginal cost, while T^1 is a lump-sum component. T^1 may be negative, that is, it may represent a payment from the supplier to the buyer. Importantly, a supplier can *condition the offered tariff to the buyer purchasing all of her requirement from it*. The first buyer decides. Then, firm E decides whether it wants to enter and if so it sinks its entry cost f and transactions take place. In the second period, active suppliers offer two-part tariffs to the second buyer: $t_j^2(q_2) = c_j q_2 + T_j^2$ with $j = I, E$ and the second buyer decides.

The incumbent and the entrant produce the same good and their marginal costs are $c_E = 0 < c_I < 1/2$. The entrant has yet to pay a fixed

entry cost f when competition for the first buyer starts, with:

$$\underline{f} \equiv \frac{1}{2} - \frac{(1-c)^2}{2} < f < 2 \left[\frac{1}{2} - \frac{(1-c)^2}{2} \right] \equiv \bar{f} \quad (2.8)$$

Assumption (2.8) implies that the demand of the second buyer alone is insufficient for the entrant to cover the entry cost, while demand of both buyers is enough. The latter condition also ensures that entry is welfare-beneficial.

Proposition 2.11 *Equilibria of the game are as follows:*

- **(Exclusion)** If either $c_I < 1 - \sqrt{2}/2$ or $c_I \geq 1 - \sqrt{2}/2$ and $f > 1 - \frac{3(1-c)^2}{2}$, then firm I and firm E offer to B_1 the tariffs $t_E^1 = f - \frac{1}{2} + \frac{(1-c)^2}{2}$ and $t_I^1 = c_I q_1 + (1-c)^2 - 1 + f$. The incumbent conditions its tariff to the exclusivity requirement. B_1 buys from the incumbent, then firm E does not enter the market and the incumbent supplies the second buyer at the tariff $t_I^2 = c_I q_2 + \frac{(1-c)^2}{2}$.
- **(Entry/Expansion)** If $c_I \geq 1 - \sqrt{2}/2$ and $f \leq 1 - \frac{3(1-c)^2}{2}$, then firm I and firm E offer to B_1 the tariffs $t_E^1 = \frac{1}{2} - (1-c)^2$ and $t_I^1 = c_I q_1 - (1-c)^2$. B_1 buys from the entrant, then firm E enters the market and the entrant supplies the second buyer at the tariff $t_E^2 = \frac{1}{2} - \frac{(1-c)^2}{2}$.

Proof. Let us start from the second period. If firm E did not enter the market then the incumbent behaves as a monopolist on the second buyer and extracts its entire surplus by setting the tariff $t_I^2 = c_I q_2 + \frac{(1-c)^2}{2}$. If, instead, the entrant pays the entry cost, then competition for the second buyer results in the incumbent offering $t_I^2 = c_I q_2$ and firm E securing the second buyer by offering the tariff $t_E^2 = \frac{1}{2} - \frac{(1-c)^2}{2}$. Such an offer allows firm E to extract from B_2 the additional surplus that its superior technology can generate relative to the incumbent's offer.

At the end of the first period, firm E decides not to enter the market if the first buyer bought from the incumbent: the rents extracted from the second buyer are insufficient to cover the entry cost, by assumption $f > \underline{f}$. If the first buyer addresses firm E , then firm E decides to pay the entry cost if the overall rents extracted from the two buyers are large enough: $\pi_E = T_E^1 + \frac{1}{2} - \frac{(1-c)^2}{2} - f \geq 0$.

In the first period the two firms compete for the first buyer. The best offer that firm E is willing to make, compatible with the break-even

constraint, is $\tilde{T}_E^1 = f - \frac{1}{2} + \frac{(1-c_I)^2}{2}$. By assumption (2.8), such offer of firm E entails a positive lump-sum payment. Is it profitable for the incumbent to match this offer? To answer this question one has to take into account that the incumbent has a higher marginal cost than firm E . Hence, when the two firms supply at their respective marginal cost, firm E generates a larger buyer's surplus. Then, in order to match the best offer of firm E , the incumbent must discount firm E 's lump-sum payment of the amount $\frac{1}{2} - \frac{(1-c_I)^2}{2}$, which is the additional surplus produced by firm E when the suppliers sell at marginal cost. Hence, in order to win the first buyer, the incumbent must offer a lump-sum payment slightly lower than $T_E^1 - \frac{1}{2} + \frac{(1-c_I)^2}{2} = f - 1 + (1 - c_I)^2$. This payment is negative, by assumption (2.8). Such an offer is profitable for the incumbent if and only if:

$$\pi_I^{TOT} = T_I^1 + T_I^2 = f - 1 + (1 - c_I)^2 + \frac{(1 - c_I)^2}{2} > 0 \Leftrightarrow f > 1 - \frac{3(1 - c_I)^2}{2}. \quad (2.9)$$

Note that if $c_I < 1 - \sqrt{2}/2$, then $1 - \frac{3(1-c_I)^2}{2}$ is below \underline{f} . Hence, the incumbent manages to exclude for any feasible value of the entry cost. If, instead, $c_I \geq 1 - \sqrt{2}/2$, then the entry cost must be sufficiently large. In both cases, the incumbent wins the first buyer, firm E does not enter and the incumbent sells also to the second buyer.

If, instead, $c_I \geq 1 - \sqrt{2}/2$ and the fixed cost is not large enough, then it is firm E that wins the first buyer by fixing a lump-sum payment which amounts to the incumbent's best offer (that is, $T_I^1 = -\frac{(1-c_I)^2}{2}$) increased by $\frac{1}{2} - \frac{(1-c_I)^2}{2}$. Hence, $T_E^1 = \frac{1}{2} - (1 - c_I)^2$. Following this offer it is profitable for firm E to pay the entry cost. Then, firm E supplies also the second buyer. ■

Note that the reason why inefficient exclusion may arise is the same as in the predation model analysed in Chapter 1, or in the model with market-share discounts analysed above. By being more efficient than the incumbent, firm E has larger total surplus to offer to buyers as compared to the incumbent, which favours firm E when competing for the first buyer. However, the incumbency advantage implies that firm E cannot extract the entire surplus from the second buyer, whereas the incumbent manages to do so. This puts firm E at disadvantage when competing for the first buyer. If the latter effect is sufficiently strong, the less efficient incumbent wins competition for the first buyer.

Lemma 2.12 *At the exclusionary equilibrium the incumbent's tariff $t_I^1 = c_I q_1 + (1 - c)^2 - 1 + f$ entails a negative lump-sum payment, that is, a compensation to the first buyer.*

Proof. $T_I^1 = (1 - c)^2 - 1 + f < 0$ by assumption (2.8). ■

The implication of this lemma is not only that the incumbent needs to suffer losses on the first buyer to exclude. Another implication is that the incumbent needs to impose an exclusivity requirement (or to condition the discount on the buyer purchasing a sufficiently large amount from it). Absent exclusivity, indeed, the buyer would have an incentive to cash in the lump-sum compensation from the incumbent and then buy from the entrant. The incumbent would not be able to exclude the rival.

As a consequence, should the incumbent be constrained not to impose the exclusivity requirement, then it would have to attract the first buyer by setting a *below-cost* linear price. This introduces an allocative distortion which limits the incumbent's profitability and makes inefficient exclusion less likely, as the following proposition shows.

Proposition 2.13 *If the incumbent cannot include an exclusivity requirement in the offer made to the first buyer, then it offers a linear tariff to the first buyers and equilibria are as follows:*

- **(Exclusion)** *If either $c_I < 2 - \sqrt{3}$ or $c_I \geq 2 - \sqrt{3}$ and $f > 2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2$, then firm I and firm E offer to B_1 the tariffs $t_E^1 = f - \frac{1}{2} + \frac{(1-c_I)^2}{2}$ and $t_I^1 = \left(1 - \sqrt{1 - 2f - c_I^2 + 2c_I}\right) q_1$, with $1 - \sqrt{1 - 2f - c_I^2 + 2c_I} < c_I$. B_1 buys from the incumbent, then firm E does not enter the market and the incumbent supplies the second buyer at the tariff $t_I^2 = c_I q_2 + \frac{(1-c_I)^2}{2}$.*
- **(Entry/Expansion)** *If $c_I \geq 2 - \sqrt{3}$ and $f \leq 2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2$, then firm I and firm E offer to B_1 the tariffs $t_E^1 = \frac{1}{2} - \frac{(1-c_I + \sqrt{3}(1-c_I))^2}{8} > 0$ and $t_I^1 = \frac{1}{2}(1 + c_I - \sqrt{3}(1 - c_I)) q_1$. B_1 buys from the entrant, then firm E enters the market and the entrant supplies the second buyer at the tariff $t_E^2 = \frac{1}{2} - \frac{(1-c_I)^2}{2}$.*

Proof. In period 1, the linear pricing that allows the incumbent to match firm E's best offer must be such that, by buying at such price, B_1 obtains

the same surplus as in the case in which it accepts firm E 's offer:

$$\frac{(1 - p_I^*)^2}{2} = \frac{1}{2} - f + \frac{1}{2} - \frac{(1 - c_I)^2}{2} = 1 - f - \frac{(1 - c)^2}{2}. \quad (2.10)$$

The above equality is satisfied if:

$$p_I^* = 1 - \sqrt{1 - 2f - c_I^2 + 2c_I} < c_I. \quad (2.11)$$

Note that such a price is below the incumbent's marginal cost. Hence, the incumbent suffers losses on the sales to B_1 . The surplus extracted from the second buyer dominates such losses, if the following condition is satisfied:

$$\pi_I^{TOT} = (p_I^* - c_I)(1 - p_I^*) + T_I^2 \quad (2.12)$$

$$= \left(1 - c - \sqrt{1 - 2f - c_I^2 + 2c_I}\right) \left(\sqrt{1 - 2f - c_I^2 + 2c_I}\right) + \frac{(1 - c_I)^2}{2} > 0. \quad (2.13)$$

The above inequality is satisfied iff:

$$f > 2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2. \quad (2.14)$$

Note that if $c_I < 2 - \sqrt{3}$, then $2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2$ is below \underline{f} . In this case, the incumbent profitably wins the first buyer and inefficient exclusion arises for any feasible value of the entry cost. If instead $c_I \geq 2 - \sqrt{3}$, then the entry cost must be large enough for exclusion to arise at the equilibrium. If $c_I \geq 2 - \sqrt{3}$ and $f \leq 2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2$, then the incumbent cannot profitably match firm E 's best offer. At the equilibrium the incumbent offers the lowest linear price compatible with the break-even condition, that is, the price that satisfies the following condition:

$$(p_I - c)(1 - p_I) + \frac{(1 - c_I)^2}{2} = 0 \Leftrightarrow p_I = \frac{1 + c_I - \sqrt{3}(1 - c_I)}{2} < c_I. \quad (2.15)$$

Firm E wins the first buyer by offering the linear tariff that gives the buyer the same surplus as the incumbent's price, that is, such that:

$$\frac{(1 - p_I)^2}{2} = \frac{1}{2} - T_E^1 \quad (2.16)$$

with $p_I = \frac{1 + c_I - \sqrt{3}(1 - c_I)}{2}$. ■

Lemma 2.14 *Exclusion is less likely as compared to the case in which the incumbent can impose an exclusivity requirement.*

Proof. The critical value of the incumbent's marginal cost below which exclusion is always feasible is larger when the exclusivity provision can be imposed as compared to the case in which the incumbent is forced to use linear pricing: $1 - \sqrt{2}/2 > 2 - \sqrt{3}$. Moreover, the critical value of the entry cost above which exclusion is feasible is lower when the exclusivity provision can be imposed as compared to the case in which the incumbent is forced to use linear pricing: $1 - \frac{3(1-c_I)^2}{2} < 2c_I - c_I^2 - \frac{\sqrt{3}}{4}(1 - c_I)^2$. This shows that in the former case exclusion is more likely to arise at the equilibrium. ■

2.5 From Theory to Practice

Most firms – not only dominant ones – resort to price discrimination in its various forms (rebates, discounts, coupons, etc.). In Section 2.2.2 we have seen that, abstracting from possible exclusionary reasons, price discrimination is not necessarily welfare-detrimental. First, the short-run (that is, for given productive efficiency) welfare effects of price discrimination are ambiguous (in particular, price discrimination may allow new customers to be reached and increase market participation). Second, there exist reasons to believe that price discrimination, by raising expected profits, may give stronger incentives to invest thereby resulting in dynamic efficiency gains. Third, a ban on price discrimination might help an input monopolist to enforce its market power and solve its commitment problems when dealing with downstream firms (for example, retailers), thus leading to higher prices. Fourth, price discrimination may create efficiency gains of various nature. For instance, a quantity discount applied to individual transactions may reflect cost savings due to scale economies in transportation, distribution, or packaging. But efficiency gains do not arise just because of cost savings: a market-share discount that applies to the entire requirement of a buyer over a given reference period may help guarantee a supplier that the buyer will make sufficient orders, inducing it to invest more in the relationship with the buyer; a two-part tariff (which represents a way to implement a quantity discount), by reducing the linear component of pricing, will tend to decrease allocative inefficiencies.

For all of these reasons, it would not make sense to suggest a blanket prohibition on price discrimination, whatever form it takes.

Yet, we have also seen in this chapter that price discrimination may be a powerful tool for an incumbent firm which would like to exclude smaller and new rivals. In particular, in industries where scale economies matter, an incumbent firm may use a 'divide-and-conquer' strategy and offer a good deal to some buyers which account for a critical volume of sales, and a very high price to the remaining buyers. Deprived of the orders of the former, the rival will not be able to compete effectively for the latter, which will guarantee the profitability of this exclusionary strategy for the incumbent.

It follows that rebates schemes offered by incumbent firms may have anti-competitive effects and may be worth investigating, especially when they offer large discounts, they are contingent on buyers buying very large volumes or requirement shares, and are individualised.

Theory of harm As also emphasised in Chapter 1, a crucial step in any investigation of abusive behaviour is the spelling out of a theory of harm and the assessment of whether the facts of the case are consistent with the proposed theory. The analysis conducted in the earlier sections can help formulate a theory of harm.

First, the theories that support an exclusionary motive for price discrimination and rebates rely crucially on the *existence of scale economies*, whether due to fixed costs, learning effects, demand externalities, two-sided market effects, or other reasons. (Recall also that price discrimination may be used not only to exclude a rival from the industry altogether, but also to relegate it to a market niche. For instance, a rival that already exists may need scale to recover investments in additional capacity, to build a more connected network, to expand in new markets, and so on. So the fact that the rival has reached minimum efficient scale in a particular segment of the market does not imply that scale economies do not play a role any longer.) This means that one should assess whether the rebates are aimed at preventing the rival from reaching minimum efficient scale (or scope) in the market or a subset of it. This implies, for instance, that if the rebates were directed to just a small proportion of the buyers, it would be difficult to argue that they could attain an exclusionary objective.

Second, there must be a strong *incumbency advantage*, proxied by a strong asymmetry in the availability of a crucial infrastructure/input, or in the pre-existing customer base, and more generally by a high and persistent market share of the incumbent; this would be reinforced by the presence of switching costs, by the infrequency of purchases and by

demand externalities.⁴⁸ As a consequence, the existence of a *dominant position* is a necessary requirement for the finding of abusive rebates also from an economic theory perspective, and the stronger the dominance the more severe the anti-competitive concern should be. On the contrary, in situations in which dominance is weak, because the dominant firm faces (relatively) strong rivals, one should be rather sceptical of a rebate scheme having an exclusionary effect.⁴⁹

Third, exclusionary price discrimination is more likely when *demand is fragmented*, that is, when there are many buyers whose individual demand is insufficient to sponsor entry and who are unable to coordinate their purchasing decisions. In contrast, *buyer concentration* may significantly alleviate anti-competitive concerns. In this respect, central purchasing agencies, by grouping orders of affiliated buyers, may create sufficient buyer concentration as to avoid inefficient exclusion. Indeed, as stressed in Section 2.3.2.2, also looser institutions that favour communication among buyers may be beneficial by mitigating the risk of coordination failures (that is, the risk of buyers ending up with a worse outcome than would have been the case if they all chose differently and bought from the incumbent's rival, for example). A careful analysis of the buyer's situation and bargaining power is therefore important, in particular to see to what extent they need to rely on the dominant firm for a part or most of their sales.⁵⁰

Fourth, exclusionary price discrimination is more likely if *downstream competition is sufficiently weak*. As also discussed in Chapter 1, when downstream competition is sufficiently intense, the demand of a single buyer may be large enough for the upstream rival to cover the entry cost or, more generally, for it to achieve efficient scale. Then, the incumbent cannot take advantage of coordination failures or of divide-and-conquer strategies to exclude a more efficient rival. Similarly, it eliminates the possibility for the incumbent to extract more rents than the rival from later buyers. In sum, the critical mechanism that rationalises inefficient exclusion does not operate any longer in the presence of fierce downstream competition. As

⁴⁸ Recall also that, as discussed in Section 2.3.2.2, where there are network effects and there exist strong asymmetries between an incumbent and entrants, *interoperability obligations* may represent an important instrument to prevent the incumbent from relying on such imbalances to exclude rivals.

⁴⁹ See the discussion of *Michelin II* in Section 2.7, for example.

⁵⁰ The discussion of the *Meritor* and *Eisai* cases (discussed in Section 2.7) offer some insights in this respect. In *Meritor*, the buyers had to rely on the defendant (Eaton) for a large part of the range of products they needed, whereas in *Eisai* it appeared they could easily switch away from the alleged infringer.

noted in Chapter 1, however, measuring the degree of competition is not straightforward, although some proxies may include the degree of product differentiation (either physical or geographical), the prevalence of capacity constraints, and in mature industries the importance of switching costs and lock-in effects.

Fifth, the analysis of the literature reviewed in Sections 2.3.2.3, 2.3.2.4 and 2.3.3 suggests that *contracts that reference rivals*, that is, contracts that condition the terms of trade on how much the buyer purchases from the incumbent's rival such as exclusivity discounts or market-share discounts – raise more severe anti-competitive concerns (all else equal) because they limit the amount of profits that the incumbent needs to sacrifice so as to exclude.

Finally, the theoretical analysis (see Section 2.3.2.2) has shown that the exclusionary effects of *individualised* discounts are stronger than those of *standardised* discounts, suggesting that they should be regarded with more suspicion when used by a dominant firm. This means, for instance, that volume rebates which are not modulated across buyers (like in the *Michelin II* case described in Section 2.7) are less likely to be exclusionary.

Which standard for below-cost pricing? There certainly exist some parallels between predatory pricing and (exclusionary) price discrimination. We have seen that in most of the models of price discrimination analysed in this chapter an incumbent achieves exclusion of an efficient rival by setting prices below costs to some buyers (or to some units sold to each buyer) and recouping any loss by charging supra-competitive prices on other buyers (or on other units). However, we have seen that – similarly to predation – there may exist circumstances where exclusion can be achieved without going below costs. For instance, we have shown that if buyers suffer from coordination failures, exclusionary outcomes may exist without the dominant firm making losses on any buyer. Moreover, other mechanisms discussed in Chapter 1 for predatory prices which do not necessarily involve below-cost pricing may also rationalise exclusionary price discrimination.⁵¹

Given the parallel with predation and given that there is no ground to prohibit price discrimination in general (we have seen in Section 2.2.2 that price discrimination may have beneficial effects for consumers, for instance

⁵¹ Imagine that there are buyers whose demand is partially captive to the dominant firm. Imagine also that the dominant firm wants to target sufficiently low prices (even though above costs) to the contestable demand of such buyers, so as to limit the profits that the rival is able to make and prevent it from obtaining external funds for investment. Market-share discounts may well serve this purpose.

by expanding the market), it may be natural to also adopt a policy rule that first assesses the *existence of a dominant position* and then prevents a dominant firm from making discounts that involve *below-cost pricing* on some groups of customers.

Like for predation, for the sake of administrability and in order to avoid the risk of chilling legitimate competition, we could then propose to adopt a safe harbour for rebate schemes which result in above-cost prices. In principle, then, one could espouse a price-cost test also for rebates cases, establishing that if the dominant firm's prices for each customer are *above* an upper bound of costs, the case should be dismissed. (For the possible cost benchmarks, see the discussion in Chapter 1, and in particular Section 1.4.3.)

Instead, if the dominant firm prices are *below* the lower bound for *some* significant portion of customers (or for some portion of the customers' demand which represents a significant part of total demand), there should be a strong presumption of abuse, with the dominant firm having the burden of proving otherwise, in particular that there are objective justifications for below cost pricing for those specific customers or those specific units.

Finally, with prices falling *between* the two thresholds, the plaintiff (or the agency) should have the burden of proving that the rebate is abusive. Similarly, the agency should bear the burden of proof when the evidence is uncertain, for instance because there are different methods of estimating cost thresholds. As emphasised in the discussion in Chapter 1 on predatory pricing, particular attention should be given to the formulation of a rigorous theory of harm, showing that there is a precise theory that rationalises the abusive conduct and that the facts of the case are consistent with that theory. Moreover, this process should also involve the analysis of the effects of the alleged abusive conduct on the market. (See Chapter 1, and in particular Section 1.4.3. We refer the reader to that section also for the discussion concerning the role of documentary evidence on the intent to exclude.)

Having pointed out the similarities, let us discuss also the differences between predatory pricing and exclusionary rebates schemes and let us make some important caveats on the possible use of a price-cost test for rebates.

First, it is important to emphasise that the incumbent's exclusionary strategy does not call for unprofitable offers on average, but offers which entail losses on some units or buyers while being recouped through high prices on remaining units or buyers. Then, differently from predation in

which the same price is charged to all the buyers that are purchasing in a given moment of time, with rebates recoupment may well be simultaneous. As a consequence, the price-cost analysis should consist of assessing whether the dominant firm suffers losses *on the sales to the critical buyers*, or on the *contestable share* of these sales, thereby checking whether the *discounted price* offered to such buyers (or for such contestable units) is below some relevant cost benchmark. The formulation of a coherent theory of harm, by allowing to understand whether and why some buyers (or some portions of buyers' demand) are critical for the success of the rival, offers some valuable guidance also in this case and avoids a mechanical implementation of the price-cost test.

Second, in a case of alleged abuse through anti-competitive discounts, an agency should verify that any below-cost sales are *significant*, in the sense that they should involve enough sales to lead to exclusion (it would be difficult to believe that 1 per cent of sales below cost may be sufficient for exclusion), and non-occasional (special offers may be justified on various grounds, for instance by the intent to dispose of stock in excess).⁵²

Third, in the case of quantity discounts, it is also important to take into account the position of the target that qualifies a buyer for the discount. If the buyer's total requirements are insufficient to reach the target, then the rebate scheme is ineffective.

Fourth, the identification of the discounted price to be compared with the appropriate cost-benchmark requires some additional care when discounts are offered in the form of retroactive rebates, as we discussed in detail in Section 2.3.2.5.

Fifth, and more important, whereas we have little doubt that a strong presumption of infringement exists whenever prices charged to a significant group of buyers are below a lower bound of costs, we would not exclude the possibility of finding anti-competitive harm for above-cost pricing in some particular cases. This is because some types of rebates – and particularly those which reference rivals, such as discounts conditional on exclusivity, or conditional on buying a very large share of the buyer's

⁵² In some cases, though, even a relatively small coverage may result in rebates being exclusionary. Consistent with this point, the Advocate General's Opinion in *Intel* stated: 'It is certainly true that thresholds may prove problematic due to the specificities of different markets and the circumstances of each individual case. For example, where loyalty rebates target customers that are of particular importance for competitors to enter or expand their share of the market, even modest market coverage can certainly result in anti-competitive foreclosure. Whether that is the case will depend on a number of factors specific thereto' (see para. 142).

requirements – have a potentially strong anti-competitive effect, and may thus deserve a more stringent treatment by competition authorities. Accordingly, it may be justified to find issues with loyalty rebates that result in above-cost pricing, provided however that the circumstances of the case are sufficiently serious – this would be the case, for instance, if the incumbent firm holds a particularly strong dominant position, and the rebates cover a significant proportion of the market – and that there exists a coherent theory of harm that explains the mechanisms which allow the dominant firm to exclude and its incentives for doing so.⁵³

Finally, whether or not the price-cost test is considered to be dispositive, there is little doubt that it is an important source of information. In particular, when the incumbent's prices are above its (appropriate measure of) costs, an as-efficient rival will have the possibility to make an attractive offer to buyers without incurring losses. This begs the question of why the rival is not successful, which in turn reminds us again of the importance of formulating a convincing theory of harm.

Similarly, we believe that any evidence of below-cost pricing should be complemented by a careful analysis of the case, which includes an understanding of the theory of harm, and in particular of whether there is a coherent strategy of exclusion.

2.6 Case-law

In this section, we briefly review the landmark cases in both the US and the EU, insofar as they significantly affected the case-law in these jurisdictions.

2.6.1 United States

In single-product rebates' cases, the US courts have followed the *Brooke* predatory standard (discussed in Chapter 1): to be successful, a plaintiff must be able to show that the alleged predator charged prices below (some notion of) cost; and there is a 'dangerous probability' that the alleged predator will recoup the entire sacrificed profits. Therefore, as Kobayashi (2005) notes, provided the volume discounts lead to prices above cost, even rebates based on market share targets or quasi-exclusivity have been deemed lawful (but see the recent *Meritor* judgment discussed below).

The US case-law on pure single-product rebates is therefore rather scant – being 'squeezed' between predatory cases and multi-product rebates ones (which we consider in Chapter 4).

⁵³ See also Fumagalli and Motta (2017a).

Concord Boat concerned boat engines. In the mid-1980s Brunswick (the main supplier of inboard and stern drive marine engines, with a 75 per cent market share) began offering discounts of up to 3 per cent to boat builders provided they sourced 60–80 per cent of their engine needs from it. The Appeals Court, overturning a jury verdict, ruled (in 2000) that Brunswick's discount program did not amount to exclusive dealing and it did not foreclose actual nor potential competitors. The Court also stressed that the loyalty programme could be terminated at very short notice and that the plaintiffs failed to show the existence of substantial barriers to entry in the boat engine market. Further, the Court dismissed the potentially predatory nature of the rebates, since no party had brought evidence that (post-rebate) prices were below cost.

In *Virgin Atlantic*, the Appeals Court (Second Circuit) affirmed a lower court's ruling that rebates by British Airways ('BA') to travel agencies (in the form of higher commissions) fell within the law. BA had a 39 per cent share of the slots at London Heathrow, its main airport hub. Its fidelity scheme offered to travel agencies and corporate customers involved the provision of rebates based on the proportion of flights bought from BA over a range of routes/flights purchased by such clients (some discounts were also based on individual growth targets). The court rejected both the theory of predatory foreclosure and the price-cost test⁵⁴ proposed by the plaintiff, on the basis that it did not provide sufficient evidence for either. This outcome was therefore very different from the European one in *Virgin/BA*, which we discuss in Section 2.6.2. (We will also discuss in Section 2.7 two analogous

⁵⁴ *Virgin Atlantic* argued that its expert evidence submitted showed that BA's incentive agreements with travel agents resulted in revenues from incremental passengers (that is, those who flew with BA only because of the agreements with travel agents) below the incremental costs of the additional flight frequencies introduced. BA's economic expert critiqued this analysis because in his view such incremental revenues should have included all revenues from those flights. The Courts agreed on this point with BA's economic expert, so the plaintiff's test was dismissed.

The Court also added that *Virgin Atlantic* had failed to prove that the incremental capacity was solely due to the incentive schemes. Ideally, we note that one would like to distinguish between the contestable and the non-contestable parts of demand, and to perform a comparison of revenues and costs of the former. However, in practice, this may prove quite challenging, in which case it would appear to us more reasonable to compare all incremental revenues from additional capacity to all incremental costs of it.

Even this approach though may not be fully satisfactory, particularly if the incumbent's incremental flights cannibalised some of the revenues from its existing flights (that is, an incumbent's passenger is choosing one of the incremental flights but would have otherwise flown with the incumbent anyway on another of its – non-incremental – flights).

cases that arose in South Africa, *SAA I and SAA II*, spelling out economic considerations more fully.)

Finally, we shall discuss in detail in Section 2.7 two recent cases – *Meritor* and *Eisai* – where central for the judgments was the question of whether loyalty rebates could be a violation of the Sherman Act even if prices were above costs. This could arise where exclusivity – and not pricing – is the predominant conduct. Interestingly, the judges arrived at different conclusions (mostly, on the basis of different factual evidence). Notably, the Supreme Court declined to review *Meritor*, implying that there might be more uncertainty on this issue for several years to come.

2.6.2 European Union

The legal and policy contexts surrounding price discrimination in Europe are very different from those of the United States. The original critical stance in Europe was due to the view that price discrimination entailed market segmentation (often country by country) and this would amount to a ‘failure’ of the Single Market project.⁵⁵ One of the landmark cases in this respect is *United Brands*, where the Court of Justice upheld the Commission’s finding that United Brands had abused its dominant position by (*inter alia*) charging different prices for its bananas across the different Member States (without corresponding cost differentials) and by including contract clauses aimed at preventing parallel imports.⁵⁶ The Commission and the courts proceeded likewise in the treatment of vertical agreements (for example, between a manufacturer and different national wholesalers) that would limit parallel imports.⁵⁷

⁵⁵ The principle of economic integration among Member States of the (now called) European Union was already enshrined in the original Treaty of Rome (1957). The ‘founding fathers’ had a core (politically driven) goal of abolishing economic and trade barriers between Member States: any price discrimination across countries would have been interpreted as evidence that this integration was not taking place. This led to the so-called Single Market mantra that is, a strong stance by the European Commission (and the Community Courts) against firms employing different practices and prices for the same goods across Member States, regardless of any possible economic justification for such behaviour. For further discussion, the reader is directed to Motta (2004: sections 1.2.2, 1.3.1.4 and 7.4.1.3) and Whish and Bailey (2015: 23–4 and 663).

⁵⁶ The same approach was taken by the courts in *BPB Industries, Tetra Pak II* and (for some of the practices) in *Irish Sugar*.

⁵⁷ For further discussion, the reader is directed to Whish and Bailey (2015: 752–3). While the legal principle would differ here, the underlying economics are essentially the same, that is, the desire for the manufacturer to discriminate prices according to different demand patterns (which can actually be welfare-improving as we saw earlier in this chapter).

In this book, however, we look at cases where the concern has been that dominant firms have used price discrimination to exclude a rival. In this sphere, there has been an evolution – at least in principle – in the role of economics as applied to the Commission’s assessment of rebates. Starting from a very formalistic approach which would sanction certain forms of rebates (as well as other practices) by their very nature, regardless of their effect, the Commission slowly began to change its stance. This resulted in the publication of the Guidance Paper on how to enforce abuse of dominance in 2009 (the ‘Guidance Paper’).⁵⁸ Yet, as we shall see both in this Section and in Section 2.7, where we look at the economics of *Intel* and *Tomra*, the application of the ‘more economic approach’ principles evoked in the Guidance Paper has not been embraced wholeheartedly in the Commission’s decisions, and the recent General Court judgment in *Intel* (see the discussion in Section 2.7.1) might signal a return to a more formalistic approach towards rebates in the EU.

The first landmark case concerning single-product rebates⁵⁹ as a potentially exclusionary practice was *Michelin I*. In the late 1970s, Dutch customers of replacement tyres were offered individualised rebates provided they exceeded some annual sales targets. The Commission ruled these to be abusive (and the Court agreed with it that these rebates were in breach of competition law). Michelin was found to have a market share of around 60 per cent (the closest competitor had an 8 per cent market share). The Commission objected to the rebate scheme mostly because discounts were individualised and not transparent.⁶⁰ In *Michelin II* (discussed in more detail in Section 2.7.7), the abusive practice was that of offering French customers of replacement tyres (over the 1980s and the 1990s) rebates on the condition that certain annual targets were met. In this second case, however, the main rebate offered was standardised (and retroactive), that is, in principle available to any customer.⁶¹ In this sense, this decision set an

⁵⁸ European Commission (2009).

⁵⁹ We summarise *Hoffmann La Roche* (a landmark case on bundled fidelity rebates) in Chapter 4.

⁶⁰ According to the Commission, the dealers did not know targets and bonuses with certainty, and Michelin could exercise discretion when interpreting them. The Commission found that this uncertainty reinforced the discriminatory and anti-competitive nature of the rebates. From an economic perspective, it is not clear why a company would not want to make its rebates policy clear, given that these rebates allegedly have an incentivising purpose: if a retailer is not sure that when buying an extra unit she will be offered a discount, in which sense the discount has an exclusionary objective?

⁶¹ The ‘progress bonus’ and the ‘individual agreements’ were actually individualised rebates. The ‘Club des amis Michelin’ scheme offered further preferential terms (including

important precedent for the Commission, stating that even standardised quantity discounts (intended as discounts on volumes purchased over a certain period) may be unlawful. The General Court agreed with the Commission, pointing out that the rebates were abusive because they did not reflect any cost savings. Confronted with the fact that Michelin's market share was falling during the period when the contested practice took place, the General Court rebutted that the scheme was abusive because, in its absence, it was 'very probable' that Michelin's market share would have fallen even further (para. 245 of its judgment); however, the General Court provided no reasoning as to why this counterfactual was the correct one.

Compagnie Maritime Belge is probably the most prominent European case on an abuse of *collective* dominance. The case concerned the use of 'fighting ships' by a shipping conference, Cewal, between Europe and the ports of (then-called) Zaire, in the late 1980s. This shipping conference was made of several different firms offering liner transport services (cargo) to shippers. There was only one independent competitor, Grimaldi and Cobelfret ('G&C'). The conference members adopted a strategy aimed at eliminating this competitor according to the Commission: depending on the schedule of the competitor's shipments, the Cewal member with the closest departure time would offer a 'fighting rebate' to the shipping company in question, matching or undercutting G&C's rate. The Commission found an infringement due to the very existence of a cartel. In addition, though, it found that the companies forming part of Cewal abused their collectively dominant position through their selective price cuts; Cewal was also found to illegally provide loyalty rebates of 12.5 per cent to buyers who accepted full exclusivity. On appeal, Cewal argued (*inter alia*) that G&C actually saw its market share increase in the relevant market over the period of the alleged abuse from 2 per cent to 25 per cent. So exclusionary effects arising from the selective price cuts against G&C allegedly failed to materialise. The General Court instead resorted to the same counterfactual argument as in *Michelin II*:

the applicants rely on the increase in G & C's market share in order to maintain that the practice complained of had no effect and hence that there was no abuse of a dominant position. The Court however considers that, where one or more undertakings in a dominant position actually implement a practice whose aim is to remove a competitor, the fact that the result sought is not achieved is not enough to avoid the practice being characterized as an abuse of a dominant position within

rebates) conditional on a certain fraction of total customer requirements being sourced from Michelin. See our discussion of the case in Section 2.7, as well as Motta (2009).

the meaning of Article [102 TFEU]. Besides, contrary to the applicants' assertions, the fact that G & C's market share increased does not mean that the practice was without any effect, given that, if the practice had not been implemented, G & C's share might have increased more significantly. (Para. 149)

However, the Court provided no reasoning behind such a conjecture. Moreover, the Court added that an abuse finding does not require there to be actual adverse effects on consumers (which, as a matter of economics, is quite unorthodox).

From a legal perspective, this case was key because the Courts upheld a Commission's finding that selective price cuts can be abusive in their own right, regardless of whether they entail below-cost pricing.⁶² We return to this case in Section 2.7, where we discuss the case alongside a model of collective predation proposed by Harrington (1989).

Besides the market segmentation rationale mentioned at the beginning of this section, the practices employed in *Irish Sugar* also had an exclusionary motive, at least according to the European Commission.⁶³ There, the dominant player in the Irish sugar market (with a market share of nearly 90 per cent) adopted a rebate strategy featuring selective price cuts (near the country's border with Northern Ireland) aimed at fending off competition from importers. As we noted in Section 2.3.1, prices were nevertheless above costs: so the observation that prices were lower in areas where the incumbent faced competition than in areas where it did not was unlikely, in our view, to be evidence of a clear exclusionary strategy.

Another landmark case was *Soda-ash – ICI*. The European Commission found that Imperial Chemical Industries ('ICI') abused its dominant position in the 1980s in the UK market for soda-ash (sodium carbonate, mostly used in glass production) by offering various forms of conditional rebates. The Commission performed no price-cost analysis: the rebates were judged abusive because the Commission considered them to have the effect of guaranteeing near-exclusivity to ICI, thus unlawfully preventing customers from switching to imported products. In this respect, the decision was quite formalistic.⁶⁴

One of the most prominent European cases on rebates is *Virgin/BA*, where the Commission found that British Airways ('BA') abused its dominant position by rewarding travel agents on a discriminatory basis

⁶² See Whish and Bailey (2015: 793). Yet, according to these authors, the case involved a number of special circumstances, implying that its precedent value may be less strong than at face value.

⁶³ The decision was mostly upheld by the General Court.

⁶⁴ The decision was mostly upheld by the General Court.

and through (retroactive) target rebates, offered via higher commissions. (We review two analogous South African cases, *SAA I* and *SAA II*, in more detail in Section 2.7.5.)

The case related to flight bookings in the early 1990s, that is, at a time where internet sales were essentially negligible: travel agents handled 85 per cent of all air travel sales in the UK. In stark contrast to the US verdict discussed in Section 2.6.1, both the General Court and the Court of Justice took the same view as the Commission. They argued that the incentive scheme put in place was particularly strong, as once the target was reached by a travel agent (equal to 95 per cent of the previous year's sales), the higher commission (beyond the base level of 7 per cent payable to an agent) would apply to *all sales* made by the agent, like in *Michelin II*. As BA had a larger sales base, in order to offer the same financial benefit to an agent in £ terms, BA's competitors would have needed to offer a much higher commission on their sales in percentage terms. Moreover, both the Commission and the Courts put a lot of emphasis on explaining that the rebates did not reflect any cost savings, which at the time appeared to be the only defence available to a dominant firm offering discounts in the EU, whether individualised or not. The Commission's decision did not refer to any price-cost test nor to losses made by BA on any of its sales. Yet the General Court reported (at para. 266 of its judgment) that the Commission – responding to BA's pleas – stated that due to the high marginal commissions offered, 'BA sold a large number of tickets at a loss'. We are not aware of what data supported this claim. In any case, the General Court referred to this in passing and did not elaborate on it.

Instead, the General Court (at para. 298 of its judgment) considered some of the effects of BA's policy and used a similar logic to the one we highlighted above in *Michelin II* and *Compagnie Maritime Belge*. It therefore noted that one should not interpret the modest growth over time in the market share of British Airways' competitors as evidence of the lack of abusive behaviour; instead, one should rather compare this with the competitive counterfactual where their growth would have been more substantial. The Commission used the same argument in its decision (see para. 107). However, neither the Commission nor the General Court provided any reasoning (let alone economic evidence) as to why that counterfactual would have been the correct one.

In 2006 the Commission issued its decision in *Tomra*. Tomra supplied reverse vending machines ('RVMs'), used in the collection and recycling of

empty drink containers.⁶⁵ At the end of the 1990s, Tomra's market share – depending on the precise market definition – was between over 80 and over 95 per cent across the European Economic Area as a whole. Tomra offered its customers (typically retail outlets and supermarkets) exclusive contracts, discounts on quantity commitments and individualised retroactive rebates. The Commission noted:

Although the agreements, arrangements and conditions found in this case contain different features such as explicit or de facto exclusivity clauses, undertakings or promises to purchase quantities corresponding to a significant proportion of the customers' requirements or retroactive rebate schemes related to the customers' requirements, or a combination of them, they all have to be seen in the context of Tomra's general policy directed at preventing market entry, market access and growth opportunities for existing and potential competitors and eventually driving them out of the market. (Para. 283 of the decision)

It was thus satisfied that 'Tomra's practices tended to restrict competition, that is to say, were clearly capable of having that effect' (para. 285). However, the Commission also decided to additionally investigate the likely effects of Tomra's practices, which it found to be restrictive of competition. The Commission additionally maintained that Tomra had failed to explain how its discounts could have adequately reflected any cost-savings (for example, based on scale economies). The General Court endorsed the Commission's overall conclusions, but clarified that the Commission was under no obligation to carry out an effects-based analysis as well (which the General Court therefore ignored). The Court of Justice confirmed the General Court's judgment in full. In particular, it clarified that the General Court's formalistic assessment of a rebate scheme offered by a dominant firm was the correct one, rather than the more economic-based followed by the Commission. We discuss this case in more detail in Section 2.7.2.

In 2009, the Commission found Intel to have abused its dominant position in the market for central processing units ('CPUs') for the x86 architecture. (See Section 2.7.1 for a detailed discussion of the case.) The Commission fined Intel a record €1.06bn for having offered abusive rebates (granted to buyers conditional on purchasing a sufficient amount of their requirements from Intel), as well as used so-called 'naked restrictions' (payments conditional on customers cancelling or delaying the launch of

⁶⁵ As the Commission explained, RVMs were originally used only for the collection of empty drink containers which could be refilled by manufacturers. Later, RVMs started collecting disposable or nonrefillable containers too. Some RVMs also identify the container just disposed by a consumer and return the deposit due to her, depending on the container type and size, as well as on local or national legislation.

new AMD-based products – AMD being Intel’s only significant competitor, according to the Commission). The Commission found that the discounts amounted to ‘fidelity rebates’, which the Court of Justice defined in *Hoffmann-La Roche* as ‘discounts conditional on the customer’s obtaining all or most of its requirements – whether the quantity of its purchases be large or small’ (para. 89 of that judgment). And it argued that according to the case-law, fidelity rebates are illegal. Although it argued it did not need to carry out an economic analysis, the Commission then performed the as-efficient competitor test, and found that in order to compensate a customer for forgoing an Intel rebate, an as efficient rival would have had to offer an effective price which was often below its own average avoidable cost. Such rebates, the Commission concluded, were therefore exclusionary.

In June 2014, the General Court upheld the Commission decision in its entirety, and in a very formalistic judgment it established that Intel’s rebates belonged to the category of exclusivity rebates, defined as those conditional on a customer purchasing most or all of its requirements from the dominant firm. These, according to the General Court, are exclusionary by their very nature, so that no economic analysis is needed to establish the illegality of such discounts. The General Court set such exclusivity rebates apart from two other categories: quantity rebates (based solely on the volumes purchased), which should generally be deemed legal; and fidelity-building rebates (which are nevertheless not directly linked to any exclusivity or quasi-exclusivity to the dominant supplier), where all circumstances would need to be assessed to determine whether such rebates are anti-competitive. Intel appealed to the Court of Justice and the judgment is still pending, at the time of writing. In the meanwhile, however, the Advocate General’s Opinion was that the General Court erred in (i) identifying a separate category of exclusivity rebates and in (ii) maintaining that, when dealing with exclusivity rebates, the Commission does not need to consider all circumstances of the case. While it is unclear at the time of writing whether the Court of Justice will follow this Opinion, such a degree of explicit support for an effects-based analysis by an Advocate General is remarkable and signals that there are very different views among EU judges on whether to follow a form-based or an economic-based approach when dealing with rebates and more generally cases of abuse of dominance. This is hardly surprising given the mixed signals on the need for effects-based analysis in the recent EU jurisprudence. At one end of the spectrum, we have seen the formalistic approach in *Intel* (General Court) and *Tomra* (General Court and Court of Justice). *Post Danmark II* (Court of Justice) appears more balanced on

this issue (see below). By contrast, *Post Danmark I* (Court of Justice in its Grand Chamber) appeared much more open to effects-based analysis (see below), and as we have just seen the Advocate General in *Intel* even more so. Similarly, the Court of Justice in *Groupement Cartes Bancaires* also stressed the need of effects-based analysis – although the allegations there concerned anti-competitive agreements under Article 101 TFEU and therefore different legal standards may apply.

As just alluded to, the Court of Justice did show more sympathy for an effects-based approach in *Post Danmark I*, decided in 2012. In 2004 *Post Danmark*, the Danish incumbent postal operator, was found by the Danish Competition Council (Konkurrenceradet) to have abused its dominant position through selective price cuts in the Danish market of unaddressed mail. In particular, the Council found that *Post Danmark* had selectively targeted three key customers (all large retailers) of its smaller rival, *Forbruger-Kontakt*. In the case of one of these customers, the prices *Post Danmark* charged were above average incremental cost but below average total cost (prices instead exceeded average total cost in the case of the other customers). The Danish Supreme Court (Højesteret), on appeal, referred the matter to the Court of Justice. In its 2012 ruling, the Court of Justice held that a dominant firm does not commit a *per se* abuse by selectively offering prices that are below its average total costs, but above the incremental cost of serving a customer. Other factors must be taken into account to establish an abuse, such as the firm's intention and the effect of the pricing policy, including the actual or potential exclusion of an as-efficient competitor. The Court of Justice also reaffirmed that objective justifications and efficiencies evidenced by the dominant firm should be taken into account.

On 6 October 2015, the European Court of Justice handed down its *Post Danmark II* ruling on referral by a Danish court. The case concerned a rebate scheme implemented again by *Post Danmark* (which at the time also had some segments of the market still reserved to it).

The rebates at issue were not loyalty rebates like in *Intel*, but 'retroactive standardised rebates', that is quantity discounts available to any buyer and awarded when purchasing a certain quantity during a period of one year.⁶⁶

The Court stated that to determine whether a retroactive standardised rebate infringes Article 102 TFEU, it is necessary to consider all the circumstances of the case, including the criteria for the granting of the

⁶⁶ The Court defines as 'quantity rebates' only those rebates which are granted in respect of an individual order; see paragraphs 27–28 of the judgment.

rebate, the extent of the dominant position of the defendant, and the conditions of competition prevailing on the market. It also specified that the as-efficient competitor test is not a necessary condition for finding that a rebate scheme is abusive, but it must be regarded as one tool among others in the assessment of the rebate. However, the Court also added that, in cases such as the one at hand, the as-efficient competitor test is of no relevance, because the dominant firm held a very large market share and because the structure of the market (including the existence of a statutory monopoly in some segments of the market) would make the emergence of an as-efficient competitor practically impossible. In other words, the Court effectively maintained that when an inefficient rival exists (and it is unlikely that it can reach efficient scale), a retroactive standardised rebate can be abusive even though it results in prices above costs.

Finally, the Court stated that in order to establish the abusive nature of the rebate scheme, its effects should be probable, but there is no need to show that the effects are serious or of appreciable nature. This implies that even if a very small proportion of the buyers were interested by the rebate, this could be enough to raise competition concerns. The Court justified this position by the ‘special responsibility’ of a dominant firm not to impair genuine, undistorted competition, and the fact that when market competition has already been weakened by the presence of a dominant firm, any further weakening of the structure of competition may constitute an abuse.

2.7 Cases

In this section we comment upon a few cases from different jurisdictions concerning various forms of rebates. We focus on exclusionary cases, and do not deal instead with pure price discrimination cases – where a firm is engaging in practices which aim at setting different prices to different customer groups in order to extract higher rents from them – although the latter may be considered incompatible with EU competition law because it may be interpreted as going against the EU’s market integration objective.

2.7.1 Intel

Intel is a recent European case that attracted a lot of attention. We first analyse the Commission’s decision, which moved some steps towards an effects-based approach, and then comment on the General Court’s

judgment, which is instead very formalistic and represents a setback for those who hoped for a more economic approach towards rebates in Europe. As mentioned in Section 2.6.2, the Advocate General's Opinion on this case disagreed with the General Court's assessment; but at the time of writing, it is not clear whether the Court of Justice will follow this Opinion once it hands down its judgment.

2.7.1.1 The Commission's decision

First, we review the basic facts of the case based on the information available in the Commission's decision. Next, we summarise the arguments that led the Commission to its infringement finding. Finally, we suggest how some of the models we presented in this chapter could help interpret the case in a more economically coherent way.

The industry and the case in a nutshell Intel was the leading global supplier of microprocessors, or central processing units ('CPUs', that is, the 'brain' of a computer) for the x86 architecture.⁶⁷

CPUs were purchased by original equipment manufacturers ('OEMs'), who then integrated them into desktops, laptops or servers. OEMs then sold these final products to end-users, either directly or via retailers. The geographic scope of the market in the decision was worldwide. Intel was found to have a market share in excess of 70 per cent over the period of the abuse, which was found to occur from October 2002 to December 2007.

AMD was deemed to have been Intel's only significant competitor in this market since 2000. The Commission noted that market participants believed that AMD's products were high-quality and innovative, and thus a viable alternative to those supplied by Intel (paras 150–64 of the decision). Thus, the Commission explained, Intel's behaviour should be assessed in the context of the growing competitive threat that AMD represented.

The industry was characterised by significant barriers to entry. These could be grouped into four categories: (i) Very large sunk costs in the form of R&D: for instance, just between 2003 and 2007, Intel spent over \$26bn on R&D (by way of comparison, Intel's total annual revenue for 2007 was about \$38bn);⁶⁸ (ii) large costs related to intellectual property: if an entrant failed to develop alternative technologies not infringing existing patents, it would have to pay Intel and/or AMD to licence their patented

⁶⁷ The CPU was the most expensive item within the total cost of a computer (the precise figure was confidential, see para. 109 of the decision); it determined by and large the computer's performance.

⁶⁸ See para. 139 of the decision, footnote 151.

technologies (see paras 129–30 and 856–8); (iii) very large fixed costs of production: ultra-clean, high-tech manufacturing facilities, called fabs, used to cost about \$2.5bn–\$3bn each, and took several years to build (see paras 859–66). These facilities were also very expensive to maintain, and this contributed towards fixed costs; (iv) product differentiation, through strong branding: this was evidenced, *inter alia*, by very high marketing costs; Intel was reported as having the fifth most valuable brand in the world, estimated at \$32bn (paras 867–74). The Commission calculated that Intel spent between 14 and 17 per cent of its revenues from x86 CPUs on marketing, over 1997–2007 (para. 868).

The nature of these facilities was also found to constitute a barrier to *expansion*, since capacity (determining output) was fixed in the short run and long lead times were needed to expand facilities (paras 116–18). At the same time, to reap scale economies, utilisation rates had to be high, roughly in the 70–100 per cent range.⁶⁹

As for the specific abuses contested, the first category analysed by the Commission concerned conditional rebates to four major OEMs. These were granted provided the OEMs sourced a minimum share of their x86 CPU requirements from Intel (80 per cent for NEC, 95 per cent for HP and 100 per cent in the case of Dell and Lenovo). Intel had also offered Media-Saturn-Holding ('MSH', Europe's largest PC retailer at the time) payments conditional on it only stocking Intel-based PCs, for over 10 years.^{70,71}

'Naked restrictions' represented instead the second abuse: Intel was found to have paid customers (OEMs) in return for cancelling or delaying the launch of new AMD-based products. This was an industry with constant innovation: Intel and AMD kept churning ever faster microprocessors, with the existing ones becoming quickly obsolete. Being able to

⁶⁹ Intel went as far as saying that incremental cost was lower than average cost over the entire reasonable range of a plant's size, meaning that it would be cost-minimising to produce at full capacity, see para. 861.

⁷⁰ Note that Intel did not sell directly to MSH, so these were simply conditional payments, not discounts on goods sold – the economics, as we shall see, are essentially the same though.

⁷¹ The Commission explained the apparent inconsistency between the period of this agreement and that of the abuse found as follows: '[The Commission] identified that the Intel conditional payments to MSH have been ongoing from October 1997 to at least 12 February 2008. However, the Commission uses its discretion not to pursue in the present Decision Intel's conduct targeted only at MSH for the periods from October 1997 to September 2002 and after December 2007'. (Para. 1640)

delay a rival's product entry even by six months would yield a significant competitive advantage over that rival.

The Commission's assessment at a high level On the legal front, the Commission classified Intel's discounts as 'fidelity rebates' and relied on the Court of Justice in *Hoffmann – La Roche* to rule them unlawful.⁷²

In terms of economics instead, it relied on a notion of 'leveraging' market power from the non-contestable to the contestable share of demand. The non-contestable share of demand was sizeable due to Intel's strong position vis-à-vis end-users, so that OEMs could not be successful without stocking Intel's products. Using the Commission's terminology:

[...] Intel is an unavoidable trading partner. The rebate therefore enables Intel to use the inelastic or 'non-contestable' share of demand of each customer, that is to say the amount that would anyhow be purchased by the customer from the dominant undertaking, as leverage to decrease the price for the elastic or 'contestable' share of demand, that is to say the amount for which the customer may prefer and be able to find substitutes. (Para. 1005)

Although the Commission did not specify the precise theoretical mechanism at work, a possible theory of harm might be based on the economic theories presented earlier in this chapter, where a dominant firm may use its incumbency advantage (where the existence of captive customers may play an important role) to exclude rivals. We shall come back to our views on a possible rationale of exclusion in this case below.

The Commission's as-efficient competitor test As for the evidence put forward to sustain the above theory, the Commission applied the as-efficient competitor ('AEC') test, which we described at an abstract level in Section 2.3.2.5. The idea behind this test was for the Commission to establish at what price a hypothetical competitor (which was as efficient as Intel) would need to offer its CPUs so as to compensate an OEM for the loss of an Intel rebate. Besides the choice (and then the actual computation) of a cost measure (average avoidable cost or AAC in this case),⁷³ this test required the analysis of two other factors, as the Commission pointed out: first, the size of the contestable share of an OEM, that is, the proportion of an OEM's requirements that could be switched to an alternative supplier; second, the relevant time horizon over which such switching should take

⁷² We discuss this case in Chapter 4, in the context of bundled rebates.

⁷³ There was a fierce debate between the Commission and Intel over which cost categories should have been deemed avoidable. We do not enter into the merit of this debate here.

place to impose a constraint on the incumbent (the Commission thought this to be one year at most).

Equivalently, the test may be implemented by estimating the ‘required share’ of an OEM that an entrant with a given unit cost must obtain in order to compete against an incumbent offering a rebate of a certain size conditional upon the customer buying a given number of units at an average sales price, and then compare it with the (estimate of the) actual contestable share. For instance, in the case of Dell, the Commission computed quarterly estimates (over the relevant period) of the minimum required share that Dell would need to switch to AMD (paying it a price equal to Intel’s AAC) in order to break even (on the assumption that Dell’s rebates from Intel on remaining units would decline by 50 per cent). The Commission found that the actual contestable share for Dell was less than the minimum required share in 9 out of the 13 quarters in question (para. 1256) and therefore that an as-efficient rival would not be viable because of Intel’s rebates. The same exercise was carried out for HP, NEC and Lenovo (as well as for retailer MSH), with similar conclusions. This, the Commission concluded, was evidence consistent with the notion that Intel’s strategy was exclusionary and abusive.

Effects on competitors and consumers Although the Commission claimed it did not actually need to show any actual adverse effects of Intel’s practices,⁷⁴ the decision did make some attempts at proving anti-competitive effects. The evidence on the actual effects of Intel’s practices is quite scant and speculative, though. Three possible channels were mentioned: (i) some notion of AMD’s exclusion or marginalisation; (ii) higher prices to end-users; and (iii) lower product quality and/or less innovation.

As for any direct effects of Intel’s behaviour on AMD, it actually emerged that AMD saw its market share *grow* (by an undisclosed amount) between 2003 and 2006, before falling in 2007 (recall that the infringement took place between October 2002 and December 2007). Without resorting to the same ‘competitive counterfactual’ argument that we already saw in Section 2.6.2 in the context of *BA/Virgin*, *Michelin II* and *Compagnie Maritime Belge* (which would have been along the lines of ‘AMD *could have performed better* absent Intel’s conduct’), the Commission nevertheless stated:

⁷⁴ The Commission justified this in various instances with the legal argument that an abuse of dominance is an objective concept, and the case-law does not require the Commission to prove any effects (see, for instance, para. 1685), either on consumers or competitors.

Naturally, it is impossible to specify what AMD's market share evolution would have been in the absence of Intel's abusive conduct. However, that evolution is consistent with the finding that Intel intensified its abusive conduct at precisely the time when AMD began to represent a greater competitive threat, and with the possibility that as a result, AMD was not able to capitalise substantially on its technological improvement during its 'window of opportunity', [Footnote: For instance, by making sufficient R&D investments to be able to also develop competitive products in the future] and has since fallen back. (Para. 1736)

As for actual direct consumer harm via supra-competitive prices, it is first important to recall that CPUs were the most expensive item of a computer. Given that OEMs were intermediate buyers, with a sufficient level of competition at that layer, one would expect a high degree of pass-through of any rebate received, thus benefiting end-consumers through lower computer prices.

In fact, rapidly falling prices were a feature of both the CPU and the computer markets, as Intel pointed out:

Microprocessor prices declined at an average rate of more than 35 percent per year during and after the [time period over which the Commission contested Intel's practices]. Microprocessor prices in fact declined considerably more rapidly than prices for personal computers, storage devices and other computer-related products. (Para. 906)

The Commission rebutted that falling prices were 'an intrinsic feature' of the industry, 'irrespective of the state of competition' (para. 908).

Clearly, this discussion of the evolution of market shares and prices uncovers the importance of building proper counterfactuals if one is to take seriously the analysis of the effects of the business practices.

As for product quality, the Commission argued that Intel's practices had the effect of reducing the incentives for AMD (and others) to invest in R&D, thus limiting consumer choice and potential benefits from further innovations:

[...] Intel's conditional rebates and payments [...] induced the loyalty of key OEMs and of a major retailer, the effects of which were complementary in that they significantly diminished competitors' ability to compete on the merits of their x86 CPUs. Intel's anticompetitive conduct thereby resulted in a significant reduction of consumer choice and in lower incentives to innovate. (Para. 1616)

By contrast, the effect of the naked restrictions – which explicitly delayed the market entry of AMD's new products – were probably a more tangible example of direct consumer harm via potentially lower product quality and more limited choice at a given point in time (see para. 1603).

Efficiency defences Intel put forward four categories of efficiencies arising from the exclusivity requirements of its rebates: (i) lower prices; (ii) scale economies; (iii) other cost savings and production efficiencies; and (iv) risk-sharing and marketing efficiencies (para. 1632).

The Commission retorted that none of the defences provided a reasonable justification for the *conditions* attached to the rebates (that is, the Commission was not objecting to the rebates in themselves, see paras 1620 and 1633, but to the exclusivity clauses).

Towards a more coherent theory of harm Our main comment on the decision is that the Commission could have tried to bring in the industry and conduct facts in a more economically coherent way. Here, we suggest that the main exclusionary theory of harm presented in this chapter may apply to the market conditions as described by the Commission in this case reasonably well.

First, Intel had an *incumbency advantage*: it had a significant portion of the demand from each of the major OEMs which was not contestable (the exact figures were marked as confidential in the Commission's decision) - that is, Intel's products had a 'must-stock' nature (see paras 1009–12) and no price offered by AMD could convince these customers to switch completely.⁷⁵

Second, *scale economies* were manifest in the form of very large, sunk R&D and marketing costs, as well as fixed costs of CPU manufacturing, which we described at the outset with our description of entry barriers.⁷⁶

⁷⁵ 'Intel's brand equity resulting from its investment in product differentiation and its installed base have given it 'must-stock' status at the OEM level, in other words, it is an unavoidable trading partner for OEMs. All the main OEMs offer predominantly or exclusively Intel-based products. Intel's must-stock status provides it with significant leverage over its OEM customers because a switch to an all- or majority-AMD product line-up would be unrealistic for them.' (Para. 870)

The following excerpt from an email between two MSH executives suggests why Intel was an essential trading partner: 'I discussed the AMD issue with [an MSH executive] and I told him, that, if [country] is not willing/able to work exclusively with Intel any more, I can exclude [country] from the contract. I asked him, if he thinks, that we would sell significantly more, and he denied. Definitely you would lose the money, and AMD is not able to compensate even part of it. Especially in the current situation (with 100% Intel you are winning a lot of market share!) it seems not very intelligent to stop this partnership now.' (Para. 698)

⁷⁶ Based on the description of the industry set out in the Commission's decision (see, for example, paras 821 and 866), the substantial scale economies in production as well as the costs associated with the development of new generations of x86 CPUs suggest that barriers (in the form of fixed costs) continued to persist even if a firm was already in the market. In any event, as we explained in Sections 2.3.2.2 and 2.5 (as well as more generally in Chapter 1), the exclusionary model based on scale economies we presented is

Third, the presence of *key customers* made exclusion easier for the incumbent. This fact was quite evident in the case:

Certain OEMs, and in particular Dell and HP, are strategically more important than other OEMs in their ability to provide a x86 CPU manufacturer access to the market. The OEMs in question can be distinguished from other OEMs on the basis of three main criteria [...]:

- market share [...];
- strong presence in the more profitable part of the market [...]; and
- ability to legitimise a new x86 CPU in the market [i.e. create consumer trust in the capabilities of a new product]. (Paras 1577 and 1584)

[...] The flip-side of large OEMs' importance in legitimising a product is that smaller OEMs are not able to do so in the same way [in particular in the corporate segment, which is the most profitable]. (Para. 1588)

According to the Commission, Intel's rebates prevented AMD from reaching these key customers, and thus from recovering the costs of its R&D (thus hindering the development of future products).⁷⁷

The Commission likewise noted that Intel's sales to HP and Dell accounted for 35 per cent of Intel's total net revenues over 2005 and 2006: these revenues exceeded AMD's *total worldwide* sales (para. 1615). But these customers were suitably compensated by Intel precisely not to switch to AMD:

[The] evidence indicates that during the period in question, Dell considered AMD to be a competitive product to that of Intel, and one which it should consider sourcing. Therefore, Dell regularly analysed the pros and cons of shifting a part of its x86 CPU requirements away from Intel to AMD.

[...] Dell invariably concluded that [Intel's] rebate, or a large part of it, would be lost if this occurred. (Paras 220–1)

These quotes depict a situation which is consistent with the models analysed in this chapter, in particular the model of Section 2.3.2.4 where buyers' demand is composed of a contestable and a non-contestable part, and by offering a sizeable discount on the contestable units, the dominant

also applicable to the marginalisation or exclusion of rivals already in the market and not just to potential entrants.

⁷⁷ 'AMD's limited access to the main OEMs is likely in itself to have had significant negative impacts on its ability to recover its research and development costs. In that respect, the Commission recalls that Dell and HP are the two largest OEMs, that they cover the entire spectrum of the market, and that each has twice as many computer sales as the next largest OEM, Lenovo.' (Para. 1615)

firm takes away from its rival the orders that the latter needs in order to achieve efficient scale.

Fourth, our analysis in Chapters 1 and 2 shows that buyer coordination would jeopardise an exclusionary strategy. In a similar vein, in Section 2.3.2.1, we summarised the Innes' and Sexton's model, where the key exclusionary mechanism is based on the incumbent offering discounts to key buyers so as to discourage them from forming a coalition against the monopolist.

Interestingly, the Decision briefly mentioned (paras 160–4) a confidential project, consisting of the attempt for leading OEMs to coordinate their purchases away from Intel and towards AMD. (Due to the confidentiality of that project, the details reported in the decision are minimal.) The project eventually failed. While we can only speculate as to the business reason behind such a failed attempt, we note that Innes and Sexton (1993) would predict that the upstream incumbent would tailor its asymmetric discounts to the buyers precisely to thwart such attempts. This 'divide and rule' strategy is evidenced at several other points of the decision. First, OEMs feared that, if they also dealt with AMD, their discounts with Intel would drop. For example, Dell estimated that its discounts with Intel would drop by 50 per cent (para. 935); or even worse, that Intel could retaliate by offering further discounts *to a competing OEM* (paras 323 and 948, for instance) thus compounding the negative effect for the original customer (OEM). In fact, it even emerged that Intel had an actual budget allocated to such discounts as a whole, as the following excerpt from an Intel internal presentation discovered by the Commission shows:

[get these Dell senior executives to] clearly understand our meet-comp process and how it applies to DELL- I.e. if they have AMD in their arsenal they'll have less meet comp exposure-hence less meet comp dollars [available] to them—even the possibility that meet-comp dollars that we're [sic] applied to DELL go somewhere else... (Cited, among several instances, at para. 1264)

Moreover, Intel could change its rebate policy very quickly and very flexibly (para. 1227).

The naked restrictions used by Intel could also be interpreted in the spirit of the Innes and Sexton (1993) model: Intel could tailor – OEM by OEM – the lump sum payment in exchange for an OEM's delay of an AMD product launch at such a level that the OEM would be better off accepting Intel's offer.

Recoupment The issue of recoupment was not discussed in the decision. Nevertheless, it is worth noting that: (i) only few large and critical buyers

were apparently offered important rebates, so one might speculate that higher prices to smaller customers could offer opportunity of recouping any losses or relatively lower margins on the large ones; (ii) all rebates were periodically renegotiated (with a quarterly frequency) so that presumably Intel may have decided to discontinue discounts as soon as it perceived that AMD's expansion projects were failing; (iii) barriers to entry (and to expansion) were considered to be very large, so it was unlikely that in the near term new entrants could discipline Intel and prevent it from increasing prices should AMD's threat have disappeared.

Some open questions This was clearly a complex case, and the possible interpretation of the facts presented in the Commission's decision in the spirit of the main exclusionary model presented in this chapter is potentially tempered by some open questions.

One relates to AMD's market share growth during the course of the infringement (the precise amount of which, as mentioned above, was not disclosed). This may suggest the lack of any actual exclusionary effects of the rebate scheme. On the other hand, under certain circumstances, a competitor ('prey') may be willing to sustain losses not to lose its customer base and perhaps while waiting for antitrust intervention. Hence, the observation of stable (or slightly growing) market shares does not necessarily signal the absence of exclusionary effects.

Another open question relates to the market coverage of the rebate scheme. The General Court pointed to evidence submitted by Intel indicating that the coverage of Intel's rebates was about 14 per cent on average, if not limited only to the contestable share of demand. In his Opinion, the Advocate General noted that while it is problematic to identify a clear threshold of market coverage above which loyalty rebates are anti-competitive,

What is certain, however, is that such market coverage cannot rule out that the rebates in question *do not* have an anticompetitive foreclosure effect. This is so even assuming that the rebates and payments in question target key customers. Quite simply, 14 per cent is inconclusive. (Para. 143 of the AG Opinion)

Further investigation may have therefore been helpful to determine whether, with such a market coverage, Intel's rebates could really foreclose an as-efficient competitor.

Final remarks Intel was a landmark case. The structural and the behavioural features of the market and its participants indicated the possibility that the incumbent had abused its dominant position.

The Commission appears to have applied the AEC test consistently with its Guidance Paper, although it could have done a better job in formulating a theory of harm in a more coherent way. As we have just explained, the case could potentially be read in light of a theory of harm based on the scale economies model emphasised in this book, although some open questions remained.⁷⁸

2.7.1.2 The General Court's judgment

As just seen, the Commission went to great lengths in this case to show that Intel's rebates failed the as-efficient competitor test. This was in accordance with the Commission's own Guidance Paper. This document, although adopted after the start of the Intel's investigation, was meant to inform the Commission's Article 102 TFEU policy on abuse of dominant position and it indicated that above-cost rebates would generally not be seen as having the potential to foreclose. However, in its decision, the Commission caveated its economic analysis with a clear statement that under the existing case-law it was not necessary to prove that Intel's rebates entailed below-cost pricing.

The 2014 General Court's judgment – which fully upheld the Commission's decision – considered such assertion by the Commission and it brushed aside the as-efficient competitor analysis of the Commission, deeming it irrelevant for cases of what the General Court considered 'exclusivity rebates'. Additionally, the General Court argued that no economic analysis was needed to establish violation, thereby (in our view) making exclusivity rebates by a firm holding a dominant position *per se* illegal.

More precisely, the General Court distinguished three types of rebates.

First, *quantity rebates*, which are linked solely to the volume of purchases and which are generally considered lawful insofar as they reflect economies of scale made by the supplier.⁷⁹

⁷⁸ In the Guidance Paper, it seems that a finding of below-cost pricing is a sufficient condition for establishing an infringement (once a firm is found dominant). In our view, instead, the price-cost test is a very important element of the analysis, but there should always be an effort to formulate a coherent theory of harm and check that the facts of the case support it (see Section 2.5).

⁷⁹ See para. 75 of the judgment. The text of the judgment, as well as its reference to a paragraph of the *Michelin II* case (a case discussed in both Sections 2.6.2 and 2.7.7) which is not illuminating, is unclear as to what types of quantity rebates are covered in this category. However, our reading of *Michelin II* is that the Court would find quantity rebates justified when they reflect efficiencies at the level of a particular transaction. We doubt that the *Intel* judges from the General Court would be ready to accept as lawful

Second, *exclusivity rebates*, namely

rebates the grant of which is conditional on the customer's obtaining all or most of its requirements from the undertaking in a dominant position. Such exclusivity rebates, when applied by an undertaking in a dominant position, are incompatible with the objective of undistorted competition within the common market, because they are not based — save in exceptional circumstances — on an economic transaction which justifies this burden or benefit but are designed to remove or restrict the purchaser's freedom to choose his sources of supply and to deny other producers access to the market [...]. Such rebates are designed, through the grant of a financial advantage, to prevent customers from obtaining their supplies from competing producers [...]. (Paras 76–7 of the judgment)

Third, *fidelity-building rebates*, namely:

other rebate systems where the grant of a financial incentive is not directly linked to a condition of exclusive or quasi-exclusive supply from the undertaking in a dominant position, but where the mechanism for granting the rebate may also have a fidelity-building effect. (Para. 78)

Whereas in the analysis of the third category it is necessary to consider all the circumstances, the General Court stated, 'the question whether an exclusivity rebate can be categorised as abusive does not depend on an analysis of the circumstances of the case aimed at establishing a potential foreclosure effect' (para. 80). This is because 'exclusivity rebates granted by an undertaking in a dominant position are by their very nature capable of restricting competition' (para. 85).

As a consequence, if a dominant firm is using an exclusivity rebate: (1) 'the Commission must only show that a practice is capable of restricting competition [whereas] it is unnecessary to undertake an analysis of the actual effects of the rebates on competition' (para. 103). (2) 'Next, given that it is not necessary to prove actual effects of the rebates, it follows necessarily from this that the Commission is also not required to prove a causal link between the practices complained of and actual effects on the market' (para. 104). (3) 'Lastly, the Court would point out that, a fortiori, the Commission is not required to prove either direct damage to consumers or a causal link between such damage and the practices at issue in the contested decision' (para. 105).

Similarly, there is no need for the Commission to demonstrate the relevance of the amount of the rebates (paras 107–9), nor the relevance of

discounts awarded by a dominant company to a buyer which has met a certain threshold of purchases in a given period of time after several transactions.

their duration (paras 110–13), nor whether only a small part of the market was concerned by the rebates at issue (paras 114–17).

There is no doubt that this is a very formalistic judgment, which – if confirmed by the Court of Justice – will likely turn back the clock of European policy towards abusive conduct: in line with the early jurisprudence of the EU Courts, there would be no need to carry out any economic analysis however simple, nor to look at the market context or the circumstances in which the rebates have been used. We therefore understand that, for a finding of abuse of dominance, it may be enough to show that a loyalty rebate has been offered, even if the firm is ‘borderline’ dominant, even if the rebate leads to prices above costs and it is offered to one customer with an irrelevant market share, and even if the rebate has no effect whatsoever on rivals and consumers.

2.7.2 Tomra

Here we discuss in more depth *Tomra*, which we have already briefly mentioned in Section 2.6.2. Tomra supplied reverse vending machines (‘RVMs’). These machines collected and recycled empty drink containers of different materials, often returning a deposit to the person inserting the empty container.⁸⁰ Tomra also provided RVM maintenance and repair services to its customers, typically supermarkets and retail outlets. The markets affected were Austria, Germany, the Netherlands, Norway and Sweden,⁸¹ and the relevant markets were defined at the country-level because of different national legislations on recycling. Tomra’s market share (by volume) was found to range from over 80 per cent to over 95 per cent across the European Economic Area as a whole by the end of the 1990s (when the alleged infringement started). Its few competitors, the Commission argued, were marginalised, taken over or forced out of the market (Prokent, the complainant before the European Commission, went bankrupt).

Tomra offered its customers three main purchase schemes: ‘exclusive contracts’ (more precisely, individualised discounts conditional on a customer sourcing most or all of its RVM requirements from Tomra); quantity commitments; and individualised retroactive rebates. Overall, about one-third of its sales were covered by such rebates according to

⁸⁰ RVMs were classified as either low-end or high-end. The former were stand-alone machines, typically used in canteens; the latter also required backroom equipment.

⁸¹ Sales in most other European markets were negligible.

Maier-Rigaud and Vaigauskaite (2006) and their reference periods typically lasted between one and three years.

The Commission's 'theory of harm' hinged upon the fact that in each country there would be a wave of RVM sales whenever (the so-called 'key years') national legislation on recycling was passed. Tomra, the Commission argued, used rebates to tie retailers in such 'key years' in order to prevent competitors from gaining a foothold in each market.

The evidence put forward by the Commission included internal documents; a thorough review of the contractual details of several agreements between Tomra and its customers; and some charts showing the effective price a competitor would need to offer to win the customer's marginal sale, which we discuss in more detail below.

The General Court's assessment was highly formalistic, and certainly a setback for those who advocated an effects-based approach in abuse of dominance cases (see Federico, 2011b, for a discussion).⁸² A vast proportion of the judgment assesses the clauses of individual contracts to determine whether they constitute exclusive dealing. It did not review the Commission's economic arguments, and it even stated that while the Commission attempted to assess the actual foreclosing effects of the allegedly abusive practice, it was not actually required to do so under the law (para. 219 of the judgment). Further, by stating that 'consumers on the foreclosed part of the market should have the opportunity to benefit from whatever degree of competition is possible on the market', and that a dominant undertaking should not 'dictate how many viable competitors will be allowed to compete' (para. 241), the General Court seems to turn back the clock to the time when neither efficiency arguments nor competition on the merits may justify aggressive business practices by a dominant firm.

Below, we discuss both the methodology followed and the conclusions reached by the Commission.

Demand was predictable, lumpy, inelastic and cyclical. The Commission explained that the introduction of national legislation meant that there were spikes in demand in so-called 'key years', and that the life-cycle of an RVM was then of 7 to 10 years, so there should have been demand spikes at national level with that cyclicalities (or if new legislation was passed). This meant that, by and large, one should not have expected substantial sales volumes during the periods in between.

⁸² The Court of Justice confirmed the General Court's judgment in full.

But if demand was predictable (that is, there was very little uncertainty) and lumpy (due to legislation), market conditions actually resembled those of a bidding market, or a procurement auction. Competition *for* the market would better fit the actual market facts. The Commission discounted this possibility rather hastily (para. 90 of the decision).⁸³ We know from Chapter 1 (and the same is true for the theories presented in this chapter) that the existence of a significant *new* demand translates into a limited incumbency advantage for the dominant firm which makes the success of exclusionary practices unlikely. To put it in the terms of the case at hand, if the market in a 'key year' is composed of – say – 80 per cent new customers and 20 per cent old customers, it is very unlikely for a firm with a majority of the share of old customers to exclude as-efficient rivals by relying on rebates or other discriminatory prices. When a large proportion of demand comes from new customers one should expect competition to be on level ground.

Entry barriers were not high. The Commission found that entry into the market was not exceedingly costly (para. 344)⁸⁴ and that there had been a number of entry attempts in the markets at issue (para. 85).

Tomra's patents over some technologies used in RVMs may have been one such barrier. Yet several entrants overcame this barrier, so it must have been feasible to produce RVMs without infringing Tomra's patents.⁸⁵

Barriers to entry were also low because of limited scale economies in production. The total size of the 'non-contestable' market amounted to 32 per cent of the total market during the 1998–2002 period, across the countries affected, according to Maier-Rigaud and Vaigauskaite (2006). But then the remaining 68 per cent of the markets taken together would likely suffice for entrants to cover their fixed costs of entry, given scale economies were limited.

Computation of the effective price a competitor should charge to match Tomra's rebates. In its decision, the Commission produced some charts indicating the effective price that a competitor should charge to match

⁸³ The Commission, in our view, conflated the lack of a formal procurement process with the lack of competition for the market. Negotiating with several parties for big, lumpy orders in 'key years' is still compatible as a matter of economics with competition for the market or a bidding market.

⁸⁴ See also Maier-Rigaud and Vaigauskaite (2006), two (then) Commission officials who wrote that 'entry was neither technically particularly difficult, nor exceedingly costly'.

⁸⁵ The Commission referred to patents only at paras 87 and 106 of the decision, stating that Tomra started alleged patent infringement proceedings only 'in a few instances' (the Commission cited two cases).

Tomra's rebates. We have already described the difficulties with such approach in Section 2.3.2.5, so we will not dwell on this here, beyond some short remarks. In principle, in this context, one should seek to determine, as robustly as possible, the contestable and the non-contestable shares of demand. However, in computing the effective price that competitors should charge to match Tomra's rebates, the Commission simply appeared to assume that all the volumes corresponding to the individual rebate *threshold* set to each customer were foreclosed to competitors.

Specifically, the Commission explained that it was irrelevant to look at *ex post*, or actual, customer demand (paras 364–90). This is because, the Commission claimed, it would have been irrational (that is, non-profit-maximising) for Tomra to set a target below a customer's realised demand (because more sales could have been extracted from that customer) or above it (since the rebate would not have been achieved, so it would have been redundant to put it in place). What mattered instead, according to the Commission, was *ex ante* demand, that is Tomra's expectation about each customer's demand – and rebate thresholds were placed exactly at that level. It was these (*ex ante*) levels, therefore, that the Commission used for its analysis of the effective price, rather than the level of demand actually observed (*ex post*).

This raised some questions, though (see also Federico, 2011b). First, Tomra showed that for several customers qualifying for its rebates, their individual demand *exceeded* the rebate threshold (that is, the target to qualify for lower prices on all units). As a result, this led to output volumes (over and above the threshold) that the Commission omitted from its calculations but that would have been potentially available to competitors (even assuming that all demand up to the threshold was not). In omitting these volumes, the calculation led to a lower effective price that rivals should match than otherwise, thus magnifying the estimate of the exclusionary effect found. Second, the Commission assumed that even in the case of customers whose actual demand was *lower* than the rebate threshold (and therefore did not qualify for it), *all* such demand was foreclosed. In principle, however, one may argue that at least *some* of that demand may have been available to competitors, particularly given that the rebate scheme had not been successful in raising the demand of those customers to the threshold levels. This assumption in the Commission's methodology therefore further reduced the estimated level of demand available to competitors and thus the estimated effective price.

In our view, the distinction between contestable and non-contestable shares of demand is important in the context of an assessment of a rebate

scheme such as in Tomra. However, the Commission's approach appears to have ignored this; instead, the Commission made some rather simplistic assumptions on the relationship between rebate thresholds and level of demand available to competitors. This approach, in our view, was not informative of whether Tomra's rebate structure could have a foreclosure effect.

Buyers appeared to be concentrated. From the decision, it appears that in each country considered there were roughly three to five major players in retailing. This may suggest relatively limited fragmentation, or even some degree of concentration. Therefore, at least potentially (we do not have sufficient information on minimum efficient scale), a single retail chain may have been sufficient to support entry (or the continued viability of an existing rival). For Tomra to prevent this, it would need to 'compensate' *all* big retail chains not to deal with a competitor, in order to foreclose it from the market. It was unlikely, therefore, that a 'divide-and-rule' strategy as that discussed in Section 2.3.2 could be successful.

In fact, not just one 'free' major retailer may have allowed sufficient scope for entry, but even smaller retailers were deemed to be important from the perspective of rivals and potential entrants, according to the Commission:

Furthermore, even smaller customers can be strategically important from the point of view of market entrants and smaller competitors. (Para. 287)

In terms of the economic model we presented, this may translate into the need for the incumbent to sufficiently compensate smaller buyers as well (and not just the major ones), making exclusion much more difficult (if not impossible) to achieve. With such market conditions, it is difficult to imagine how it could be feasible and profitable for Tomra to exclude competitors.

2.7.3 Meritor

This case concerned the heavy-duty ('HD') truck transmission market in North America. We next set out a summary of the key facts of this case based on the information available in the Appeals Court's judgment.

The market had been dominated by Eaton from 1950 to 1989, when Meritor entered and achieved, by 1999, a market share of 17 per cent. At that point Meritor decided to form a joint venture with a German company, ZF, with the aim of introducing a new (two-pedal automated

mechanical) transmission into the North American market by 2001. That type of transmission was used exclusively in Europe at that moment. In 2000 Eaton entered new long-term agreements ('LTAs') – lasting five years – with each of the *four* manufacturers of HD trucks (original equipment manufacturers, or 'OEMs'), the direct purchasers of HD transmissions.

The LTAs included rebates conditioned on the OEM purchasing at least a percentage of its requirements from Eaton. The percentage varied across OEMs and ranged from 68 per cent to 95 per cent. In some cases Eaton also agreed to make an upfront payment to the OEM entering the agreement. Each LTA also required the OEM to grant preferential treatment for Eaton in the OEM's *data book*⁸⁶ and, in two cases, to remove Meritor's products from the data book.

Following these facts, by 2003 Meritor's market share fell below 8 per cent. The joint venture with ZF was dissolved and, by the end of 2005, Meritor's share dropped further to 4 per cent. In January 2007 Meritor exited the business.

Meritor filed suit against Eaton and, in 2009, the US District Court found that Eaton's LTAs foreclosed a substantial part of the market, thereby harming competition. Eaton filed a renewed motion for judgment or for a new trial. Its claim was that Meritor had failed to establish that Eaton's conduct was anti-competitive because in all the time relevant to the case Eaton never priced below its costs. The case reached the US Court of Appeals for the Third Circuit which upheld the District Court judgment that Eaton's LTAs were anti-competitive.

The argument of the Appeal Court was that in this case price was not the vehicle of exclusion. As a consequence, evidence that prices were above costs did not make the arrangements lawful.

The Court explained that:

[...] this is not a case in which the defendant's low price was the clear driving force behind the customer's compliance with purchase targets, and the customers were free to walk away if a competitor offered a better price. (Page 37 of the judgment)

The reason was that:

[...] compliance with the market penetration targets was mandatory because failing to meet such targets would jeopardize the OEMs' relationship with the dominant manufacturer of transmissions. (Page 38)

⁸⁶ Truck buyers have the ability to select many of the components used in their trucks, including the transmissions, from the OEM catalogue, denoted as *data book*.

As one OEM executive testified, if the targets were not met the OEMs ‘would have a big risk of cancellation of the contract, price increases and shortage of supply’ (page 37). And breaking the relationship with Eaton was a risk that OEMs could not afford: the reason is that Meritor did not sell the complete range of transmissions needed by OEMs which made Eaton an unavoidable trading partners:

Eaton was a monopolist in the HD transmissions market, and even if an OEM decided to forgo the rebates and purchase a significant portion of its requirements from another supplier, there would still have been a significant demand from truck buyers for Eaton products. Therefore, losing Eaton as a supplier was not an option. (Page 37)

In other words, the fact that not all demand was contestable combined with the threat not to supply the OEMs was central to the finding that Eaton’s conduct did not rely mainly on pricing.⁸⁷

As a consequence, the Court decided to adopt a *rule of reason* approach and assessed the anti-competitive effect of Eaton’s conduct by looking, *inter alia*, at whether the extent of Eaton’s market power was significant, at the existence of high entry barriers in the HD transmission market, at the rival’s evolution of market share, at the duration of the agreements, at their coverage, and at the extent to which they could be terminated, as well as at other provisions contained in the agreements. Moreover, the Court considered whether there were possible pro-competitive justifications for the agreements.

The Court found that Eaton’s strong market power, reinforced by high barriers to entry, allowed it to coerce OEMs into accepting LTAs, and after analysing the evolution of the market it was satisfied that there was sufficient evidence that LTAs foreclosed a substantial share of the market for an extended period; and that exclusionary conduct caused antitrust injury, that is, Meritor’s inability to grow.

The decision was not unanimous. In his dissenting opinion Judge Greenberg argued, *inter alia*, that the Supreme Court’s

unwavering adherence to the general principle that above-cost pricing practices are not anticompetitive and its justifications for that position lead me to conclude

⁸⁷ Note that the mere threat to stop or limit supply is not credible: once the targets are not met the dominant firm has all the incentive to supply the non-contestable part of the demand. Hence, such all-or-nothing clauses facilitate a dominant firm in inducing buyers to accept exclusivity provided that the firm managed to build a reputation of enforcing the threat. The existence of such a reputation is a condition that should be carefully checked. See also Chapter 3 for a discussion of all-or-nothing clauses, as well as of *Dentsply*.

that this principle is a cornerstone of antitrust jurisprudence that applies regardless of whether the plaintiff focuses its claim on the price or non-price aspects of the defendant's pricing program. (Page 20 of the Dissenting opinion)

Similar criticism was also echoed in an 'Amici curiae' brief by a number of US competition experts who (unsuccessfully) urged the Supreme Court to reverse the Third Circuit's judgment,⁸⁸ whereas other authors have defended the idea that loyalty discounts should not be submitted to the same standard as predatory pricing cases and should be considered instead as similar to exclusive dealing.⁸⁹

Judge Greenberg argued that prices were indeed central to the LTAs:

Eaton's prices were the crux of the rebate program and an inextricable element of the LTAs.

[...] LTAs themselves would not exist without the reduced prices that Eaton offered as an incentive for the OEMs to enter the agreements. (Pages 40–42 of the Dissenting opinion)

Low prices, then, attracted buyers into LTAs whereas, according to the dissenting judge there was:

[...] absence of evidence in the record suggesting that Eaton would have refused to supply transmissions to the OEMs had the OEMs failed to meet the LTA's market-share targets or that Eaton at any point coerced the OEMs into entering the LTAs or meeting the targets. (Page 22 of the Dissenting opinion)

The conclusion of Judge Greenberg was, therefore, that Meritor's decreasing market share and ultimate exit from the market were due to business failures – such as inability to develop a full range of products, a key source of competitive advantage in this market, and inability to offer appealing price conditions – rather than to Eaton's practices.

Comments We believe that the facts of the case lend themselves to an interpretation along the lines of the theory of harm we have sketched out in this chapter, particularly in Section 2.3.2.4: buyers' demand is composed of a contestable and a non-contestable part, and by offering a sizeable

⁸⁸ Daniel A. Crane et al. 'Brief for eighteen scholars as Amici Curiae in support of petitioner', 28 March 2013.

⁸⁹ See Joshua D. Wright (2013) 'Simple but Wrong or Complex but More Accurate? The Case for an Exclusive Dealing-Based Approach to Evaluating Loyalty Discounts'. Speech delivered at the Bates White 10th Annual Antitrust Conference, Washington, DC on 3 June 2013. See also the following heated debate in the blog www.truthonthemarket.org, containing a number of interesting contributions by authors such as Crane, Lambert, Salop and others.

discount on the contestable units, the dominant firm may take away from its rival the orders that the latter needs in order to achieve profitability. As for other ingredients of the theory, Eaton clearly enjoyed a very strong incumbency advantage, it does seem that scale economies mattered a lot, the industry was a mature one, and the four buyers were apparently not coordinating their purchases, nor was evidence presented to the effect that any had enough market share to be able to sponsor entry alone.

The model also predicts that the dominant firm should suffer losses on the sales of the contestable units. In this respect, we do not know how the price-cost test was implemented in this case. Since Meritor was not able to offer the whole range of products, but could contest Eaton only in a portion of the transmission system market, the right test to be conducted would have been the price-cost test we described in Section 2.3.2.5. If this test was not performed, then the finding that Eaton's average price was above cost (whatever measure of cost was used) considering all contestable and non-contestable output is not informative at all. As we have repeatedly seen in this chapter, price discrimination in its various shapes allows a dominant firm to exclude by incurring losses on some units or buyers (for example, the contestable ones) which are outweighed by gains on other units or buyers (for example, the non-contestable ones). Hence, considering *all* the output of the company, the practice may well be profitable and prices above costs on average.

Suppose though that the test had been carried out properly and revealed that an efficient rival could replicate on the contestable segments of the market Eaton's effective price (namely, Eaton's price calculated once considered the rebates, see equation (2.2) in Section 2.3.2.5), so that price was indeed above (say) average avoidable costs even for the relevant contestable units. Would this be enough to dismiss the case, as Judge Greenberg suggested, or could one still find an infringement of the law, as the majority judges found?

As we pointed out in Section 2.5, market-share and exclusivity discounts are practices that reference rivals, which have the potential to cause severe anti-competitive effects and may therefore deserve particular attention by competition authorities and courts. Accordingly, it may be justified to sanction them even if they do not involve below-cost pricing. At the very least, though, we feel that a competition authority or court should rely on strong evidence to find an infringement in situations where the dominant firm is shown not to charge below costs on any significant portion of its demand. For instance, the incumbent firm should hold a particularly strong dominant position, the rebates should cover a significant proportion

of the market, and so on. Moreover, a well articulated theory of harm should be proposed, so as to explain why the rival does not manage to make an attractive offer and secure enough buyers when the incumbent's prices are above costs. In other words, a coherent story explaining the mechanism which allows the incumbent to exclude would be particularly compelling.

Comparing Meritor and Intel If both *Meritor* and *Intel* (reviewed in Section 2.7.1) share the idea that a price-cost test is not necessary to establish violation of the law in the case of loyalty rebates, the European's General Court judges in *Intel* go much further than their US counterpart, and basically argue that economic evidence is irrelevant in evaluating loyalty rebates, because they are inherently restricting competition if used by a dominant firm.

Instead, the US judges in *Meritor* (as well as commentators, see for example Wright, 2013) do not disregard at all economic considerations and, as described above, analyse whether there is evidence of anti-competitive conduct and effects.

Further, the US judges in *Meritor* do not seem to embrace a formalistic approach as they argue:

Moreover, a plaintiff characterization of its claim as an exclusive dealing claim does not take the price-cost test off the table. Indeed, contracts in which discounts are linked to purchase (volume or market share) targets are frequently challenged as de facto exclusive dealing arrangements on the grounds that the discounts induce customers to deal exclusively with the firm offering the rebate. However, when price is the clearly predominant mechanism of exclusion, the price-cost test tells us that, so long as the price is above-cost, the pro-competitive justifications for, and the benefits of, lowering prices far outweigh any potential anticompetitive effects. (Page 32 of the judgment)

Finally, one point raised by Judge Greenberg, the dissenting judge, is interesting for our discussion. He argues that even in cases in which the non-price aspects of a conduct are the main drivers of the anti-competitive effect, the price-cost test may still provide useful information, and 'would operate only as one element, though a significant one, of a court's and jury's inquiry under the rule of reason' (page 41 of the Dissenting opinion).

2.7.4 Eisai

In this section we briefly discuss *Eisai* (already mentioned in Section 2.6.1). Interestingly, the reasoning in the District Court's judgment in *Eisai* is remarkably similar to the Appeal Court's judgment in *Meritor* (discussed

just above), but the District Court's judge in *Eisai* reached a very different conclusion. As we shall see below, the Appeals Court – while affirming summary judgment in favour of the defendant (thus upholding the lower court's *Eisai* judgment) clarified the type of relevant test that should apply in similar cases.⁹⁰

This case concerned a market for drugs – in particular, the relevant market was defined as the market for 'brand-name anticoagulants'. Sanofi's Lovenox was found to have between 81.5 and 92.3 per cent of the market in the relevant period (2005–2010), and the other main drugs in the market were deemed to be Eisai's Fragmin, as well as Glaxo's Arixtra and LEO Pharma's Innohep. Under the scrutiny of the District Court's judge was the Lovenox Program, which consisted of a system of discounts based on both volumes and market shares. Similar to *Meritor*, here it was also undisputed that Sanofi's price (after discounts) was well above costs:⁹¹ indeed, it was estimated to be 17.7 times its costs.

However, in this case, the District Court's judge found in favour of the defendant: she maintained that – unlike in *Meritor* – in this case pricing was central to the practice under scrutiny, and hence the fact that price was above cost was sufficient to dismiss the case.

Crucial to this case were the accessory clauses: in particular, the District Court's judge stressed that the contracts between the hospitals and Sanofi were terminable on both sides at 30 days' notice (in *Meritor*, Eaton had the right to terminate the contract if the targets were not met); and that hospitals were required not to advantage rivals, rather than discriminate against them (in *Meritor*, the OEMs were obliged to give preference to Eaton in their data book, or even removing rivals from it, and Eaton's products were to be priced below rivals').

Given that in her view *Meritor* did not apply, and that pricing was central, the finding of price being above costs was enough to dismiss the case. However, the judge found that even under a rule of reason approach (that is, even if – like for an exclusive dealing case – the price-cost test was not dispositive) her finding would still be one of non-violation. She based this on a number of reasons, including: there was no customers' testimony of foreclosure (in other words, no evidence that any buyer had modified its purchase decision due to the programme); the market shares

⁹⁰ Interestingly, the Court of New Jersey belongs to the same Third Circuit as the Court of Appeal of *Meritor*.

⁹¹ Even in this case, though, it is not clear what exactly the price-cost test consisted of, nor what cost threshold was considered.

of Fragmin and Arixtra had increased over the period (implying absence of effects on rivals); and hospitals could ‘walk away’ from Lovenox without consequences (in other words, the aggressiveness of Sanofi’s salesforce and their threats were not credible).

Interestingly, the District Court’s judgment did acknowledge that a portion of the demand could have not been contested by the smaller Eisai, but it argued that Eisai could have competed much more fiercely:

Any alleged incontestable demand did not prevent Eisai from reducing its 85% profit margins to [...] increase its market share. [...] It may well be that Sanofi had even larger profit margins. In 2009, for example, Sanofi was able to charge a price that was 17.7 times higher than its costs, while Eisai charged 7.8 times its cost [...]. But the antitrust laws do not protect businesses from the loss of profits due to vigorous competition. (Pages 64–5 of the judgment)

These considerations echo Judge Greenberg’s arguments in *Meritor* that Meritor was not an as-efficient competitor to Eaton and that a finding of monopolisation in circumstances in which the dominant firm charges in excess of its costs would protect rivals which would not deserve it, and hence reduce competition in the market.

The Appeals Court (Third Circuit) affirmed summary judgment in favour of the defendant thus dismissing Eisai’s claims of an antitrust violation. Specifically, the Court held that Eisai had not presented evidence of substantial foreclosure of competition.⁹² Further, the Court pointed out that Eisai presented no evidence that Sanofi’s conduct caused or was likely to cause anti-competitive effects (increase in prices, reductions in output or consumer choice) in the relevant market.

However, very importantly, the Appeals Court disagreed with Sanofi’s – and with the District Court’s – argument that because Sanofi’s discounts did not lead to below-cost prices, the discounts could not be considered anti-competitive. The Appeals Court did therefore reaffirm the principle set out in *Meritor* that when pricing conduct predominates over other means of exclusivity, a price-cost test should apply. However, in relation to the specific facts of the case, the Appeals Court held that Eisai’s bundling of contestable and non-contestable demand for Lovenox acted as the primary exclusionary tool.

⁹² The Court stated that ‘identification of a few dozen hospitals out of almost 6,000 in the United States’ which allegedly wanted to buy Eisai’s drugs but did not as a result of Sanofi’s conduct did not amount to ‘substantial foreclosure’ (pages 15–16 of the judgment).

2.7.5 South African Airways, I and II

We summarised the different fate of British Airways' discriminatory commissions to travel agents across the Atlantic in Sections 2.6.1 and 2.6.2. There have also been, however, two analogous cases in South Africa, involving the national incumbent, namely South African Airways ('SAA'). We present them below, noting that the South African judges (unlike the European ones) made an effort to use economic analysis.

In *SAA I*, the South African Competition Tribunal found that SAA illegally granted incentive payments and commissions to travel agents on domestic air markets between 1999 and 2001. Some of the payments offered (so-called 'overrides') were retroactive and applicable whenever an individualised growth target was met. In addition, SAA offered incremental commissions, payable only on sales beyond a certain target (these rebates were individualised too). Moreover, through the 'Explorer Scheme', SAA incentivised individual travel agencies' *employees* by granting a free international ticket conditional upon a sales' target being met.

Let us summarise the Tribunal's analysis of these schemes.

Dominance and relevance of the incentive schemes As for dominance, SAA was found to have a share between 65 and 70 per cent of the domestic airline market. (BA/Comair had about 25 per cent of the market, and Nationwide about 6 per cent.) Under South African competition law, there was an unrebuttable presumption of dominance if a firm had a market share above 45 per cent, and in any case based on the evidence set out in the Tribunal's judgment there were no features of the market which may suggest that SAA's market power was lower than its market shares indicated.

SAA sold about 85 per cent of its tickets via travel agents (direct sales from airlines being the main alternative sales channel, since at the time internet sales were irrelevant), suggesting that practices which affected travel agents' behaviour were not without consequences in the product market. Further, travel agents appeared to have some power in steering consumers towards SAA flights: due to constant volatility in airfares, and limited information on ticket availability (recall again that these were times the internet was underdeveloped), even relatively well-informed passengers could actually end up trusting an agent when she claimed that the best option available was SAA (even in cases where it was not). It was also noted that it would not have been financially viable for existing and potential competitors to set up their own retail agencies. The Tribunal also argued

that rival airlines could not compete with SAA in their payment offers, due to their much more limited flight offerings.⁹³

Anti-competitive effects of the schemes As for the anti-competitive effect of SAA's incentive schemes, and absent empirical evidence about the effects on consumers, the Tribunal focused on the foreclosing effect of the incentive schemes.

First of all, it showed that a significant portion of the travel agents' market was subject to override agreements (the judgment mentions that all four major travel agent groups as well as a number of smaller agents, were covered by agreements). Next, the Tribunal observed that both Nationwide and BA/Comair experienced a decline in market shares, even if it acknowledged that other events (including competition on the merits by SAA) may have contributed to the decline of rivals' shares. Overall, though, it was satisfied that the more aggressive nature of the override schemes adopted by SAA had inhibited rivals from expanding while at the same time reinforcing the dominant position of SAA.

Analysis of possible pro-competitive effects If a firm is engaging in exclusionary acts, South African competition law explicitly allows for an efficiency defence, the alleged abusive practice being prohibited 'unless the firm concerned can show technological, efficiency or other pro-competitive gains which outweigh the anti-competitive effects of its act' (Section 8 of the Competition Act of South Africa).

In general, rebate schemes used by a producer to incentivise its retailers may have a number of pro-competitive effects, the main one probably being the fact that a reduction in the wholesale price paid by the retailer will likely translate into more competitive pricing downstream. However, the Tribunal ruled out this effect in the SAA case. Indeed, SAA was retaining decisions over the prices of the airline tickets, so travel agents would not have the possibility to reduce prices as a way to increase their volume of sales. Therefore, any gain they would make would not be passed on to final consumers, who would not be likely to see any price decrease.

In fact, the Tribunal argued that the payment scheme might have the effect of *increasing* final prices and biasing consumers' choice in an inefficient way. The lack of transparency on airline price tickets (customers could not compare price offers and their availability) gave travel agents the

⁹³ Federico and Regibeau (2012) formalise a game which well explains this point. See our discussion below.

possibility to distort consumers away of their preferred choice, redirecting them to SAA flights which gave agents higher commissions.⁹⁴

The Tribunal also considered but discarded the attempts to invoke efficiency arguments by SAA's representatives and consultants, and finally concluded that SAA's incentive schemes were in breach of competition law.

In *SAA II*, the Tribunal once again ruled against SAA's incentive schemes offered to travel agents, in this case between 2001 and 2005.⁹⁵ The nature of the commissions paid was essentially the same as that in *SAA I*. The Explorer Scheme had been dropped but the override payments remained broadly similar. SAA also introduced the so-called 'TRUST' agreements, which involved lump-sum payments from SAA to a travel agent conditional upon it reaching a certain sales or market-share target. This payment too was individualised and took place 'with all major travel agents' (para. 33 of the judgment).

Overall, the overrides and the TRUST agreements had a high coverage: between 70 and 90 per cent of total airlines' sales through travel agents were covered by such agreements with SAA.

Over the period in question, SAA's market share of sales of domestic air travel through agents ranged between 74 and 79 per cent. Compared with *SAA I*, there had been a few market developments, most notably the launch of low-cost carrier Kulula (owned by Comair, a full-service carrier competing within South Africa), as well as the growth of other distribution channels such as the internet and direct sales (meaning call centres or corporate agreements).

However, these factual differences did not have a material impact on the type of economic and legal analysis undertaken in *SAA II* with respect

⁹⁴ The Tribunal wrote: 'It is also clear, that at least during the reference period, the level of internet usage for ticket bookings was still in its infancy, and hence the best tool for making the market more transparent to consumers was not yet fully functional. Consumers are also aware that ticket prices are a function of demand over time. Since consumers have no access to cycles in demand from time to time, informational asymmetries are easily maintained. Consumers need not suspect they are being duped simply because they hear different ticket quotes for a similar service at different times. They may assume the difference is simply a function of demand movements over time.

Furthermore, ticket availability is also never a constant. Were agents to suggest that a rival airline had no tickets available, particularly if they are SAA's rivals who are known to have less capacity, consumers may have no reason to suspect that this may not be so and given that occupancy, like price, is so transient a notion, it would also not alert suspicions if the consumer later became aware that seats were available. In essence, the consumer's ability to police opportunist behaviour by agents, is seriously constrained in this type of market, because of the informational asymmetries that exist' (paras 208–9).

⁹⁵ See Federico (2013) for a detailed analysis of this case.

to SAA I.⁹⁶ Travel agents still constituted the main distribution channel (ranging from over 60 per cent of Nationwide's sales to over 85 per cent of SAA's) and were therefore again the sole focus of the analysis. The Tribunal thus reached the same conclusions as in SAA I and ruled against SAA.

One difficulty in dealing with cases such as *BA/Virgin* and SAA I and II is to identify the likely theory of harm behind the rebates schemes and the test that one should use to see whether the schemes are anti-competitive. Indeed, the rationale behind exclusionary discrimination seen in this chapter hinges upon 'divide-and-conquer' strategies in markets characterised by scale economies of some sort. Scale economies are certainly important in the airline sector, but the models we have presented so far are unlikely to offer a convincing explanation of the exclusionary potential of the rebates used by, for instance, SAA. Travel agents are not retailers which *buy* a product from the manufacturer to resell it to final consumers, as in the models seen in Section 2.3.2.2. Rather, they are intermediaries which offer a service to the airlines in exchange of a commission, but without ever owning the product and without having the power to set prices, which are set by the airlines.⁹⁷

Federico and Régibeau (2012) offer a model which addresses these issues. They show that a firm with a wider range of products (and higher market share) tends to outcompete a smaller rival when bidding for agents' favouritism, and that this is more likely when firms can resort to fixed payments (interpreted as fidelity rebates) rather than linear incentive schemes, that is, payments (or discounts) on a per-ticket basis which do not vary with the volume of tickets sold. (Intuitively, when payments are fixed, a firm with higher market share has an advantage over smaller rivals.) In their model, anti-competitive effects from rebates arise because agents engage in *directional selling* (that is, exploiting a lack of sufficient levels of information by consumers, for example around the alternatives available, they induce consumers to buy a product which is not their preferred one).

⁹⁶ In SAA II, though, the judges relied more on empirical evidence brought by the parties. For instance, the assessment of the effects of SAA's incentive schemes included the comparison of the evolution of sales of travel agents which had not participated in the schemes – where SAA lost share – with those which did participate – where SAA increased its weight.

⁹⁷ Note that commissions are interpreted as rebates even though they are payments made by airline companies to intermediaries, and not discounts (or compensations) made by airline companies to the buyers of airline tickets.

The model is quite convincing in showing that in industries characterised by information asymmetries between consumers and agents, incentive schemes aimed at the fidelisation of agents may be anti-competitive. But in general, incentive schemes may also encourage agents to engage in sales-enhancing activities which also benefit consumers, so one would like to understand under which conditions rebate schemes are pro- or anti-competitive.⁹⁸

2.7.6 *Compagnie Maritime Belge*

We already mentioned the *Compagnie Maritime Belge* case in Section 2.6.2. In brief, Cewal, a shipping conference comprising several firms offering cargo services to shippers in Western Africa (holding over 90 per cent of the market), sustained a cartel in the 1980s by fighting the only rival (G&C) through selective price cuts. Depending on the departure time (day) of G&C's shipment, the Cewal member with the nearest departure time would offer a 'fighting rebate' to the customer in question, matching or undercutting G&C's rate. Yet even after rebates, prices in *Compagnie Maritime Belge* were actually above cost, that is, they failed a traditional predatory test. Moreover, as we noted in Section 2.6.2, Cewal's selective price cuts (together with its fidelity rebates conditional on full exclusivity) did not have any material exclusionary effect: G&C's market share in fact increased from 2 per cent to 25 per cent over the relevant period.

This case cannot be interpreted on the ground of the scale economies model discussed in this chapter. Rather, it might be rationalised referring to a model of collective predation put forward by Harrington (1989). The major result stemming from that paper was that collusion could occur even in the absence of entry barriers, due to the implicit agreement that, should an entrant join the market, this would be met with collective predation against it. Importantly, in the model, predation would only occur upon entry of a new competitor. But this, by definition, would not occur on the equilibrium path: the collusive equilibrium strategy was defined precisely as one which *deters* entry, for given parameter values. Thus, such a collusive strategy would guarantee both internal stability (no cartel member would defect) and external stability (no entrant would enter the market, anticipating the predatory response, in spite of arbitrarily small entry costs).

⁹⁸ *Comair v South African Airways* (a judgement on a consolidated damages action following *SAA II*) is consistent with the discussion of *SAA II* as set out in this section.

In fact, in *Compagnie Maritime Belge*, G&C *did* enter the market. In a strict sense, the model would then suggest that the collusive strategy was not successful, since the external stability property of the equilibrium failed; that is, the potential entrant was not deterred. Still, one could imagine an incomplete information version of the model in that paper where entry and predation may be observed at equilibrium.

Moreover, the predation against G&C was not particularly fierce. In spite of the seemingly aggressive strategies put in place by Cewal (with selective price cuts, fidelity rebates and re-scheduling aimed at attacking G&C), G&C actually saw its market share rapidly increase during the ‘predatory’ phase. The General Court rebutted that G&C would have grown even more, absent Cewal’s predatory strategy (para. 149 of the judgment).

Thus, while we still find this case relevant from the point of view of economic theory and competition policy, as it featured selective price cuts in a context without single-firm dominance, we find it difficult to see this as a convincing case of abuse of collective dominance. The fact that selective price cuts still left prices above cost, together with the entrant significantly growing its market share during the predatory phase, were clear signs that anti-competitive foreclosure did not take place.

2.7.7 Michelin II

Michelin II – which we summarised in Section 2.6.2 – concerned rebates in the French market for new replacement tyres for trucks and buses.⁹⁹ The basic rebate offered was a *standardised* volume rebate, available (and transparent) to all dealers. This basic system entailed between 18 and 54 different targets (steps), and consisted of *retroactive* rebates: once the following threshold was met, a higher discount was applicable not just on the incremental units, but on all units previously bought within a one-year reference period.

The other main rebate scheme in *Michelin II* was the so-called ‘Club des amis Michelin’, which would grant preferential terms to a dealer (including a financial contribution) provided the dealer met certain criteria, including a minimum share of total requirements sourced from Michelin (something akin to a market-share discount).¹⁰⁰

⁹⁹ In what follows, we also draw from Motta (2009).

¹⁰⁰ The case also featured some (target) rebates in the form of a ‘progress bonus’ and an ‘individual agreement’. There was also a ‘PRO agreement’ (which had a tying element) and a ‘service bonus’, which would entail a discount based on dealers offering a certain number of services to end-customers.

The Commission assessed this case in a very formalistic way, without any effects-based analysis. First, it discounted any efficiency rationale associated with the rebate schemes simply by stating that these were not reflecting transaction-specific savings. Second, it did not perform a price-cost test. Third, it did not put forward a coherent theory of harm.

Rather, the Commission stressed some supposed ‘unfairness’ in Michelin’s practice, based on the following observations. The dealer segment was very competitive (in 1995, for example, the Commission reported that the average operating margin of a dealer was 3.7 per cent of turnover, para. 206 of the decision). But rebates were paid back by Michelin two months after the end of the reference period. So, the Commission reasoned, dealers were put under a lot of pressure towards the end of the reference period, as failure to achieve the desired target would have entailed losing the rebates and thus making an overall annual loss; and, as a by-product, competitors would have had to sell at very low prices in order to match Michelin’s marginal prices on the last volumes within a reference period.

The above arguments were also espoused by the General Court, which made an additional, rather economically unorthodox argument. During the appeal process, Michelin argued that during the allegedly abusive practices (which had been in place for at least 19 years), its market share had steadily been falling and it sought to use this evidence to support its argument that it no longer had a dominant position. The General Court rebutted that the abusive rebate schemes prevented that market share from dropping even further (see para. 245 of the judgment). It gave no evidence for such a statement. If this logic were to be taken seriously, then *any action whatsoever* performed by a dominant firm could be portrayed as being abusive: it would suffice for a court to claim that competitors would be stronger absent that behaviour, without the need to substantiate such a claim.

Michelin’s competitors were not small firms but rather strong international players. While their actual market share in France was not high (Michelin’s share was in excess of 50 per cent), their collective share had been steadily rising over time, in spite of Michelin’s commercial practices having been in place for around two decades. It is hard to reconcile these market facts with exclusionary behaviour by a dominant firm.

As for the strongly competitive conditions in the dealers’ segment (that is, the downstream market), these were rather an indicator that Michelin’s ‘loyalty rebates’ were unlikely to generate exclusion. As discussed in Chapter 1 (in particular in Section 1.2.3.2) – but the results obtained for the model of predation can be reproduced for the model

of exclusionary discrimination under scale economies presented in this chapter, Section 2.3.2.2 – all else being equal, stronger downstream competition makes a dealer keener on ‘deviating’ from an exclusivity agreement with an upstream incumbent. This is because even a small competitive advantage over other dealers could translate into big market share gains (and especially so in the case of homogeneous products); this makes exclusive agreements (or loyalty-inducing discounts) less likely to foreclose.

Further, there was no indication that Michelin was selling below cost; nor that it was targeting ‘key’ dealers so as to deprive Michelin’s competitors of critical scale economies and force them out of the market. These elements could have been part of a cogent theory of harm, as we have seen in this chapter, but there was no evidence that they played a role in *Michelin II*.

Another shortfall of the decision from an economic perspective was that it did not take potential efficiency defences into serious consideration (beyond quickly ruling out transaction savings and scale economies): in fact, some rebate categories employed by Michelin involved the commitment to certain service and training levels by dealers, which would likely benefit end-consumers.

One last remark: if Michelin truly intended to exclude its competitors, why would it have stocked and sold competitors’ products at its own distribution outlets?

Exclusive Dealing

3.1 Introduction

This chapter deals with exclusive dealing, that is, contracts between a seller and a buyer whereby the buyer *commits* not to make any purchase from a competing seller.¹ We therefore distinguish exclusive dealing from exclusivity rebates (that is, market-share discounts conditional on the buyer purchasing all or most of its requirement from a given supplier) because the latter do not include a commitment on the side of the buyer, but just an offer from the seller to award a certain discount if a certain condition has been fulfilled (such discounts are discussed in Chapter 2).

As we shall see in this chapter, there exists a large body of economic literature which shows the possible anti-competitive effects of exclusive dealing, so it would be hard nowadays to deny the potentially exclusionary role of this practice. However, we find it instructive to adopt a historical approach, and start our analysis from the Chicago School arguments, according to which exclusive dealing cannot have an anti-competitive nature, to then explain why such arguments are not robust.

Until the early 1960s, US courts expressed hostility towards exclusive dealing. (We present a historical overview of the landmark cases of exclusive dealing in both the US and the EU in Section 3.6.) The basic underlying argument seemed straightforward and compelling. Suppose there is an incumbent producer, one buyer of that product, and a potential entrant

¹ Other contractual forms may effectively amount to exclusive dealing even if they do not explicitly prohibit trade with alternative sellers. Consider, for instance, long-term contracts which include a commitment on minimum volume purchases. If the minimum volume corresponds to the entire (or most of the) buyer's requirement, such contracts are *de facto* equivalent to exclusive dealing. These types of contracts were under scrutiny in the *Distrigaz* and *EDF* cases that will be discussed in Section 3.7.

which is more efficient than the incumbent. If the incumbent and the buyer can sign an exclusive contract before entry takes place, efficient entry will be deterred. Since the contract commits the buyer to purchase exclusively from the incumbent, the entrant would not be able to sell if it entered the market. Therefore, the entrant would be better off staying out of the market (thus saving entry costs); and this generates an adverse effect on consumer welfare.

This attractively simple view was later criticised by Chicago School scholars. Posner (1976) and Bork (1978), for instance, argued that using exclusive dealing to deter efficient entry would require the payment of too high a compensation to the buyer, and would therefore not be a *profitable* strategy for a seller. (The Chicago School argument will be illustrated more extensively in Section 3.2.1.) Hence, the reason why firms are willing to agree on exclusivity is that such contracts generate efficiency gains which benefit the parties involved and total welfare. This implies that competition authorities and courts should not be concerned about the use of exclusive contracts.

The Chicago School's critique had the merit of pointing out that the harsh approach adopted by US courts was not supported by sound economic reasoning. Further, it attracted attention on the potential efficiency-enhancing justifications for exclusive contracts, which had been largely overlooked previously. As such, it was enormously influential and led to a more lenient treatment of these contracts in the US. Indeed, until very recent years, we are unable to find cases in the US where the use of exclusive contracts was found in breach of antitrust law.

The Chicago School's argument, however, relied on a number of specific assumptions.² The theoretical literature developed since the mid-1980s has shown that, by relaxing such assumptions, one can formally establish that exclusive contracts can be a profitable strategy to deter efficient entry. Section 3.2 will review these contributions and spell out the circumstances under which exclusive contracts are more likely to generate anti-competitive effects. The mechanism underlying many of these papers is also the one which is the *leitmotiv* of this book: in the presence of scale economies and multiple buyers, exclusive contracts can serve anti-competitive goals precisely because they allow an incumbent firm to attract key buyers, thereby depriving the entrant of the scale it needs to operate successfully. Hence, we will devote particular attention to this class of models (Section 3.2.3). Section 3.3 will then review the literature

² See also Farrell (2005) and Whinston (2006).

which has investigated the efficiency-enhancing role of exclusive dealing contracts. Technical Section 3.4 will formalise the key models we discuss in the two preceding sections. The policy implications that can be drawn from both strands of the literature are discussed in Section 3.5, together with workable criteria that can guide antitrust intervention in this area. Section 3.6 discusses landmark decisions by competition authorities and key case-law on exclusive dealing in the US and in the EU. Finally, Section 3.7 reviews from an economic perspective a number of antitrust cases involving exclusive contracts.

3.2 Anti-competitive ED

In this section we first illustrate, in Section 3.2.1, the critique of the Chicago scholars arguing that a firm with market power cannot use exclusive contracts to exclude a more efficient rival in a profitable way. We will then offer an informal overview of the post-Chicago School theoretical models that have shown instead that, under certain conditions, a firm with market power is able to exclude an efficient competitor by engaging in exclusive dealing; and that this behaviour is profitable. Specifically, Section 3.2.2 focuses on the role of the degree of upstream competition, Section 3.2.3 on economies of scale and the effects of buyers' decisions on each other, Section 3.2.4 on the possibility for exclusivity clauses to be breached or renegotiated and Section 3.2.5 on a scenario where the incumbent and a rival compete for the ability to sign exclusive contracts with buyers. Finally, Section 3.2.6 presents some empirical evidence.

3.2.1 The Chicago School Critique

The essence of the 'Chicago School critique' is that a rational buyer would want to be compensated by an incumbent seller for accepting an exclusive deal which limits her option to address alternative sellers. More precisely, a rational buyer should anticipate that under exclusivity the incumbent would be her monopoly supplier and will set prices at the monopoly level; instead, absent exclusivity the buyer would benefit from competition between upstream firms and from the resulting drop in prices. However, due to the monopoly deadweight loss (an inefficiency arising as a result of the monopoly price level and measuring the value that is not appropriated by either buyer or seller), the incumbent's gain from deterring entry (that is, the monopoly profit) is lower than the buyer's loss from accepting a lack of

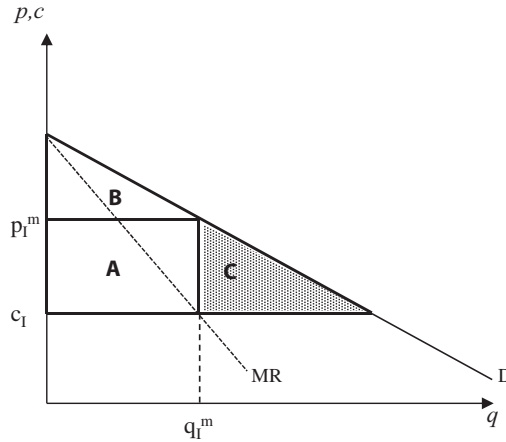


Figure 3.1. Chicago critique.

competition. This means that the incumbent cannot profitably compensate the buyer to accept an exclusivity clause.

This argument is illustrated in Figure 3.1. (See Section 3.4.1 for a formal treatment.) If an exclusive contract is signed, the incumbent sets the price p_I^m and makes a profit π_I^m , corresponding to area A. The buyer enjoys the surplus $CS(p_I^m)$, corresponding to area B. Area C depicts the monopoly deadweight loss. If no exclusive contract is signed, competition by the more efficient entrant makes the price drop to the level of the incumbent's marginal cost, c_I (assuming intense Bertrand – or perfect – competition between upstream suppliers). The incumbent's profit is 0 in this case and the buyer enjoys a surplus $CS(c_I)$ corresponding to area A + B + C. Therefore, the buyer's loss from accepting exclusivity, $CS(c_I) - CS(p_I^m)$, corresponds to area A + C and, as it is clear from the picture, is larger than the incumbent's gain from having the contract signed π_I^m (area A). It follows that the incumbent cannot induce the buyer to accept an exclusive deal (for instance using a side-payment) and make positive profits.³

³ The argument does not rely on the fact that the contractual parties know for sure that entry will take place. One would obtain the same conclusion under uncertainty about future entry, provided the incumbent and the buyer assign similar probabilities to future entry. The argument does not rely on the assumption that firms use linear prices either. Under two-part tariffs, the incumbent would extract the entire buyer's surplus $CS(c_I)$ under exclusive dealing. Hence, the incumbent's gain from exclusivity would coincide with the buyer's loss from accepting it, and it would not be *strictly* profitable for the incumbent to induce the buyer to accept an exclusive deal.

The implication of the above argument is that, if we observe firms engaging in exclusive dealing agreements, it cannot be for their anti-competitive effect but rather for legitimate efficiency-enhancing reasons (such as to protect relation-specific investments, which we discuss in Section 3.4.6 in some detail), that benefit the parties involved as well as final consumers. Using the words of Judge Robert Bork (1978):

A seller who wants exclusivity must give the buyer something for it. If he gives a lower price, the reason must be that the seller expected the arrangement to create efficiencies that justify the lower price.

The conclusion of the Chicago School critique is that there is no reason why competition authorities should be concerned about exclusive contracts.

This theory, however, depends on a number of simplifying assumptions. Once we relax them, anti-competitive exclusion from exclusive dealing may emerge. Our main objective in what follows will be to examine how changes in the main assumptions of the theory just set out will affect the results. Section 3.2.2 shows that if instead of assuming perfect competition among sellers one assumes weak enough upstream competition, it is profitable to use exclusive dealing to exclude a more efficient entrant. Section 3.2.3 relaxes the assumption that there is a unique buyer, and assumes instead that there are several (un-coordinated) buyers and that the incumbent's rival needs to secure a certain number of them in order to be profitable. This is the main theoretical part of the chapter and builds on the scale economies mechanism that features across all chapters of this book. Since theories of exclusive dealing typically assume that once the buyer has accepted the exclusivity clause, this clause is binding and will be complied without exceptions, Section 3.2.4 discusses what happens when the exclusive contract includes a provision which specifies *ex ante* a payment of damages in case the exclusivity is breached upon, as well as the case in which the exclusive contract can be *ex post* renegotiated (for instance, because all parties realise that the joint surplus would be higher if the entrant, rather than the incumbent, supplied the product). Another typical feature of exclusive dealing models is the assumption that the dominant incumbent has a first-mover advantage and can contract with buyers at a time when the entrant cannot approach buyers (for instance, because it has not entered the market yet). Section 3.2.5 shows that when the incumbent and the rival can compete for the exclusivity of the buyers, exclusive dealing can have an exclusionary effect but this is less likely to occur (and more costly for the incumbent).

3.2.2 Upstream Competition

The Chicago School's argument we sketched out in the previous section assumes that, if upstream suppliers compete (upon the entry of a rival which is more or as efficient), they do it very fiercely, so that prices are pushed down to the incumbent's marginal cost. The buyer then expects to pay a very low price if she rejects the exclusive contract, thereby requiring a large compensation from the incumbent to sign. However, imagine that upstream competition is weaker than in the above example. Then, the post-entry equilibrium price would be higher than the incumbent's marginal cost, making it easier to nudge the buyer to accept an exclusivity provision. This may create the scope for the incumbent to profitably use exclusive dealing to deter entry.⁴

To understand this intuition even better, consider the extreme case where upstream competition is the weakest. Specifically, assume that, following entry, upstream firms collude and set the monopoly price p_I^m . In this case, the buyer would expect to pay the same price irrespective of her decision to commit to exclusivity. Thus, as a matter of theory (and absent antitrust intervention on the cartel!), she would be willing to sign the exclusive contract even without any compensation. In turn, the incumbent would be willing to offer some compensation to have the exclusive contract signed: absent entry, the incumbent would earn the entire monopoly profit, whereas if entry took place, it would have to share the monopoly profit with the entrant. It follows that the exclusive contract is (profitably) offered by the incumbent and accepted by the buyer, so that entry is deterred. Of course, this example refers to an extreme situation,⁵ but it conveys the general point that the price the buyer expects to pay absent exclusivity is crucial for the extent to which the incumbent can or cannot profitably induce her to accept it. The weaker competition if entry takes place, the higher the expected price following entry and the more likely that exclusive dealing can be used to deter efficient entry.⁶

⁴ This argument is due to Farrell (2005).

⁵ Indeed, if firms collude on the post-entry equilibrium path, deterring entry does not decrease welfare since prices do not change.

⁶ Fumagalli et al. (2009) present a model which can be interpreted in the light of the reasoning above. They analyse the case where a merger between the incumbent and the entrant might take place after the decision on the exclusive dealing contract has been taken. They show that the merger may occur irrespective of whether the exclusive dealing contract has been signed, because it allows the incumbent to produce with the more efficient technology, thereby increasing industry profits. The buyer anticipates that she ends up

Another way to illustrate this point is to assume, as in Yong (1996), that the entrant is more efficient but capacity-constrained. Hence, in case of entry, the buyer cannot satisfy all of her demand through the entrant (at a relatively low price) and must also buy from the incumbent (at a relatively high price). If the entrant's capacity constraints make the resulting average price relatively high, this would lead to a low surplus for the buyer and thus a low compensation required to accept exclusivity. In a similar vein, Farrell (2005) assumes that upstream firms compete in quantities rather than in price, and shows that the incumbent can elicit the buyer's acceptance of exclusivity in a profitable way (see Section 3.4.2 for a formal treatment of this mechanism).

The intensity of upstream competition is therefore an important determinant of a possible anti-competitive effect of exclusive contracts.

3.2.3 Scale Economies and Externalities Among Buyers

Another crucial assumption for the Chicago School argument is the existence of a *single* buyer. Imagine instead that there are several buyers. If the entrant (in order to cover the fixed costs necessary to start its activities) needs to supply *more* than one buyer, then a buyer's decision to accept the exclusive contract makes it more difficult for the entrant to achieve its minimum viable scale, thereby exerting a negative externality on the other buyers (the next section will explore this point more thoroughly). The main insight behind the recent literature on the anti-competitive effects of exclusive dealing is precisely that, by exploiting this externality, the incumbent can profitably exclude efficient entry.⁷

paying the same price irrespective of her decision on exclusivity, and is willing to accept the deal behind a zero compensation. Hence, the fact that the merger occurs post-entry, thereby removing competition, allows the incumbent to elicit the buyer's acceptance in a profitable way. They also show that exclusive dealing can still be welfare detrimental, even though the possibility to merge allows the more efficient technology to find its way into the industry and makes foreclosure less of a concern.

⁷ Segal (1999) provides a general framework to study contracting with externalities, that is, situations where each agent's utility depends not only on his own trade with the principal, but also on other agents' trades. (See also Segal, 2003.) The paper shows that, when the bilateral contracts offered by the principal are publicly observed, inefficient outcomes may arise due to the principal's incentive to reduce the agents' reservation utilities. Contracting externalities arise also in the models developed in Chapters 1 and 2 to study predation (see Section 1.3) and conditional rebates (see Section 2.4). A variant of that model will be studied in Section 3.2.5 and in Section 3.4.5 of this chapter where we explore the scope

We next illustrate this idea in Section 3.2.3.1, building on Rasmusen et al. (1991), who first had the insight that externalities among buyers can be exploited to exclude,⁸ and Segal and Whinston (2000a), who refined and extended their analysis (Section 3.4.3.1 will provide a formal treatment). This discussion will treat buyers as final consumers. In Section 3.2.3.2, instead, we consider what happens when the buyers are actually firms that compete in a final market; that is, we discuss the role of downstream competition for the potential of exclusive dealing to profitably exclude efficient entry.

3.2.3.1 Naked exclusion

Suppose that the incumbent can offer an exclusive contract to *two* buyers.⁹ The contract is very simple: it includes a lump-sum compensation in exchange for the commitment not to buy from other suppliers during a given period of time.¹⁰ After observing buyers' decisions on exclusivity, a more efficient potential entrant decides whether to enter, which involves sinking a fixed cost. (Note that the incumbent is endowed with a first-mover advantage, in that it can approach buyers and have exclusive contracts signed at a time when the entrant has not materialised yet, and cannot make offers to the buyers.) The new entrant is assumed to be more efficient than the incumbent if it supplies both buyers. Then, if it enters the market and supplies both buyers at a price equal to the incumbent's marginal cost, post-entry profits cover its entry cost. However, if the entrant supplies a *single* buyer, then post-entry profits are assumed to fall short of its entry cost.¹¹

In this setting, by signing the exclusive contract, one buyer exerts a strong negative externality on the other: entry is discouraged and the free buyer (that is, the one who has not signed the exclusivity contract) loses

for anti-competitive exclusive contracts when both the incumbent and the rival can offer contracts to buyers.

⁸ In their paper exclusive dealing has the very purpose of excluding the rival. Hence, the paper has the title 'Naked exclusion' to distinguish it from the preceding paper by Aghion and Bolton (1987) – see Section 3.2.4.1 – where exclusive dealing aims at extracting rent from the entrant, with exclusion being a possible 'unintended' consequence of the practice.

⁹ We refer to the case of two buyers because it simplifies the exposition. All the results extend to the general case where there are N buyers.

¹⁰ The price is established later, when transactions take place. Assuming that the exclusive contract also commits to a price would not alter the results.

¹¹ Note that, despite the presence of two buyers only, the model captures the notion of fragmented demand-side through the assumption that no buyer alone is sufficient to attract entry.

the possibility to purchase from the more efficient producer even though she did not undertake any exclusivity obligation with the incumbent. This externality can be exploited by the incumbent to a different extent depending on the way contracts are offered, so we next consider different cases.

Simultaneous and uniform offers Let us start by considering the case where the incumbent approaches buyers simultaneously and must offer the same contractual conditions to both of them, namely the same compensation in exchange for exclusivity. Buyers simultaneously decide whether to accept or reject the offer and cannot openly coordinate their actions together. In this environment there exist both ‘entry equilibria’, where both buyers reject the contract and buy from the entrant; and ‘exclusionary equilibria’, where both buyers sign the exclusive contracts and entry is deterred; these latter equilibria are characterised by different levels of compensation offered by the incumbent (on this, see the formal analysis in Section 3.4.3.1).

To see the intuition as to why both types of equilibria exist, start by positing that the incumbent offers to the two buyers a small compensation to agree on the exclusivity clause. A situation where both buyers accept such a contract, and entry does not follow, is a possible equilibrium. Why do buyers accept exclusivity given that the modest compensation is insufficient to offset the loss suffered by paying the monopoly price instead of the competitive price? (This was the Chicago School critique we outlined earlier.) The answer is that buyers suffer from coordination failures. If a buyer *expects the other to accept exclusivity*, then she has an incentive to do the same and accept. Her order alone is insufficient to make entry profitable. Hence, she would have to buy from the incumbent at the monopoly price irrespective of whether she rejects or accepts exclusivity; but in the latter case, she obtains the compensation. Note that the incumbent extracts the monopoly profits from all buyers behind the payment of fairly small compensations: deterring entry through exclusive contracts is thus a profitable strategy.

However, the situation where every buyer rejects the above contract and efficient entry takes place is also a possible equilibrium in this setting. The reason is that, if a buyer *expects the other to reject exclusivity*, then she has an incentive to reject as well. In this case, where buyers’ expectations are ‘virtuous’, the incumbent should *fully* compensate a buyer in order to elicit acceptance on exclusivity, that is, it should offer at least the compensation that makes her indifferent between paying the monopoly price and the

competitive price. Since contractual terms cannot be discriminated by assumption in this version of the model, the incumbent should fully compensate *both* buyers; but offering such a deal is not profitable (for the Chicago School argument).

In this environment, buyers' coordination failures are key for exclusive contracts to be signed. But as economic theory cannot help us predict which equilibrium is likelier to arise, it is particularly instructive to look at experimental evidence. This literature suggests that coordination failures are not a theoretical curiosum or a red herring. In particular, Landeo and Spier (2009) specified an experimental setting that is consistent with the crucial assumptions of the theory delineated in this section. This study documented that, when contractual terms could not be discriminated, exclusion rates were fairly high and amounted to 92 per cent.¹² Landeo and Spier (2009) also allowed in a different variant of the experiment for pre-play communication among buyers, that is, allowed buyers to state their intentions before deciding on exclusivity. Although buyers could not stipulate binding agreements and remove coordination failures altogether, pre-play communication was expected to alleviate this problem and to facilitate convergence of buyers' decisions on rejection of the exclusivity clauses, thereby decreasing exclusion rates. The experimental evidence confirmed this prediction, but it is interesting to note that exclusion rates – though lower than in the absence of pre-play communication – remained relatively high (43 per cent).

Smith (2011) performed an experiment allowing for a larger number of buyers. In the baseline model (with five buyers and an assumption that signing 40 per cent of them into exclusivity would be required for exclusion to be possible), the exclusion rate was 56 per cent. This exclusion rate did not vary if the number of buyers increased to ten. Note that this is consistent with the prediction of the models we present in this section (and in technical Section 3.4.3): when no single buyer is sufficient to sponsor entry, increasing the level of buyer fragmentation does not have an appreciable impact on the likelihood of exclusion. By contrast, allowing buyers to engage in (non-binding) communication caused the exclusion rate to decrease in a statistically significant way, to 40 per cent (which supports the idea that communication can reduce the incidence of mis-coordination among buyers). Similarly, under a scenario where 80 per cent of buyers needed to sign contract exclusivity for exclusion

¹² The exclusion rate was defined as the percentage of total groups taking part in the experiment with one or both buyers accepting exclusivity.

to be possible (as opposed to 40 per cent in the baseline model) the exclusion rate fell to 13 per cent. This last result, as the author noted, supports the idea that greater antitrust scrutiny may be warranted in the presence of exclusive dealing ‘in industries with larger opportunity or fixed costs, smaller technological advantages of the potential rival, larger scale efficiencies’, which is consistent with one of the key ideas of this chapter.

Simultaneous and discriminatory offers The possibility to discriminate contractual conditions across buyers facilitates exclusion because it allows the incumbent to target the compensation at few, specific, buyers while extracting the monopoly profits from all of the buyers. This gives the incumbent a more powerful instrument to exploit the negative externality that a buyer exerts on the others by accepting the exclusive contract, thereby facilitating exclusion.

In particular, in such an environment ‘entry equilibria’ may *fail* to exist. To see the intuition imagine that, following the incumbent’s contractual offers, both buyers reject exclusivity and entry takes place. As long as twice the monopoly profits exceed the *full* compensation for a *single* buyer (a relatively weak assumption), this situation cannot represent an equilibrium: the incumbent would have an incentive to deviate by compensating *one* buyer fully (or slightly more than this) and by offering no compensation to the other; the buyer that is fully compensated would accept exclusivity irrespective of the choice she expects from the other; once this buyer accepts, the remaining buyer anticipates that her demand is not sufficient to attract entry and cannot do better than accepting the exclusive contract, even for free. Hence, the incumbent ends up extracting the monopoly profits from both buyers behind the payment of full compensation to a *single* buyer. Given the above condition, such a deviation from a candidate ‘entry equilibrium’ is profitable for the incumbent. Then there exist *only* ‘exclusionary equilibria’ in this environment.

Also in this case, the ‘exclusionary equilibria’ differ in the level of the compensation offered by the incumbent – depending on the severity of coordination failures. Under the most favourable conditions to the incumbent, exclusion is achieved following zero (or modest) compensations. However, it is worth noting that, contrary to the case of no discrimination, coordination failures are not necessary for exclusion to arise. The incumbent manages to exclude even if buyers are able to coordinate their decisions, by offering full compensation to one buyer and nothing to the other. This *divide-and-conquer* offer leads to a unique continuation equilibrium where both buyers sign the exclusive contract

and efficient entry is deterred.¹³ (See also Chapter 2, and specifically Sections 2.3.2 and 2.4, for a discussion of price discrimination offers with a divide-and-conquer purpose.)

The experimental study by Landeo and Spier (2009) partially confirms the theoretical predictions. This study finds that the ability of the incumbent to discriminate facilitates exclusion, but only when buyers can engage in pre-play communication.¹⁴ In such a case exclusion rates are found to increase from 43 to 79 per cent. Also, in line with the theory, Landeo and Spier (2009) show that under pre-play communication sellers anticipate easier coordination between buyers and are more likely to choose a *divide-and-conquer* strategy where they compensate copiously one buyer and offer a low (zero) compensation to the other.

Sequential offers The most favourable conditions for exclusion to arise occur when buyers are approached sequentially. In this case, there exists *only one equilibrium*: the incumbent will exclude at *no cost*, that is, without paying any compensation to buyers. The intuition is that, should the first buyer reject the exclusive contract, then it would be profitable for the incumbent to fully compensate the second buyer and induce her to accept. This would happen because, as discussed earlier, the demand of one free buyer is insufficient to attract entry. Hence, in spite of paying full compensation to the second buyer, the incumbent would make monopoly profits on sales to both buyers. Then, the first buyer anticipates that entry will never take place, irrespective of her decision, and is willing to sign, even for free. This induces the second buyer to do the same.¹⁵ Sequential offers represent, therefore, the most effective instrument for the incumbent to play the buyers off against each other and to profitably use exclusive contracts to exclude a more efficient rival (this element will also be very apparent where we discuss some real world cases in Section 3.7).

Contracts with staggered expiry dates In a context in which (long-term) exclusive dealing contracts with the incumbent are already in place, the fact that expiry dates are staggered gives rise to a situation in which the incumbent can approach buyers sequentially and, therefore, in which the

¹³ Instead, if the monopoly profits extracted from both buyers does not exceed the full compensation to a single buyer, we are back to the homogeneous case where both ‘entry equilibria’ and ‘exclusionary equilibria’ exist. The latter rely on coordination failures.

¹⁴ This result may be due to the fact that, absent pre-play communication, exclusion rates when contractual offers are homogeneous are already extremely high – namely 92 per cent.

¹⁵ If, instead, twice the monopoly profits are lower than the full compensation for a single buyer, there exists a unique ‘entry equilibrium’.

conditions for exclusion are the most favourable. Cabral (2014) shows that staggered expiry dates favour exclusion also when the rival has the possibility to make offers at the moment of contract renewal. Under per-period scale economies, the rival is assumed to be less efficient than the incumbent in supplying the ‘free’ buyers (whose contract is being renewed) in any given period. The reason is that the incumbent can exploit the sales to the buyers whose contract has not expired yet to achieve efficient scale in that period, while the rival cannot. The advantage provided by the interaction between scale economies and staggered contracts may be strong enough to dominate the rival’s (overall) efficiency advantage and allow the incumbent to always win competition for the contracts to be renewed.

Buyers’ ex ante commitment We mentioned at the beginning of this chapter that a distinctive feature of exclusive dealing contracts is that they include a commitment on the side of the buyer not to purchase from competing suppliers. To see the relevance of *ex ante* commitment imagine that, as in the models described above, the incumbent makes the offers first, before the entrant shows up, but buyers do not need to commit to buying from the incumbent at that point. Imagine that the incumbent offers fairly small compensations, as in the exclusionary equilibrium with simultaneous and homogeneous offers that we discussed above. If buyers do not decide at that point, then the entrant would enter the market. It would anticipate that, once it has sunk the entry cost, it would secure all the buyers by (slightly) undercutting any offer that is profitable for the incumbent; and that by so doing it would cover the entry cost. This is because the entrant is more efficient than the incumbent, and covers the entry cost when supplying all the buyers at a price equal to the incumbent’s marginal cost. *A fortiori*, it would cover the entry cost by supplying all the buyers at a price (or slightly below a price) that is profitable for the incumbent.

Similarly, consider the case in which the incumbent adopts a divide-and-conquer strategy and offers a generous compensation to one buyer and nothing to the other. If buyers do not commit *ex ante* and the deal is overall profitable for the incumbent – because the monopoly profit extracted from the two buyers exceeds the ‘full’ compensation paid to a single buyer – then the more efficient rival can enter the market and make a profitable counter-offer. By so doing, it would secure all the buyers and it would cover the entry cost.

Then, absent *ex ante* commitment, in order to discourage entry the incumbent should make an initial offer that the rival cannot profitably

match. However, since the rival is more efficient, such an offer would be unprofitable for the incumbent. The general point is that lack of *ex ante* commitment allows buyers to decide when all the offers are on the table, which facilitates the rival in making profitable counteroffers and makes it impossible for the incumbent to exploit its first-mover advantage so as to deter efficient entry. The same argument applies to a situation in which there is no first-mover advantage and both the incumbent and the rival initially offer long-term contracts to buyers, with transactions taking place in later periods. (We will discuss that model in Section 3.2.5.) Also in that case, in the absence of *ex ante* commitment on the side of the buyer, the incumbent would not manage to exclude the more efficient rival.

For this reason, the anti-competitive concern of exclusive dealing contracts is more severe than the one of exclusivity (or loyalty) discounts.¹⁶ Indeed, a crucial difference between an offer which promises to offer a discount upon verifying *ex post* that the buyer has bought only from the incumbent (a so-called ‘exclusivity discount’) and an exclusive dealing contract resides in the *ex ante* commitment nature of the latter.

Minimum share requirements Minimum share requirements (‘MSRs’) are contracts where a buyer commits to source a minimum proportion of its requirement from a certain seller (exclusive dealing being the limit case where the buyer commits to purchase all its requirements from the same seller).¹⁷

Since under MSRs a smaller share of a buyer’s requirement is exclusive relative to fully exclusive contracts, MSRs are often regarded as a weaker version of fully exclusive contracts, and as such the presumption may be that MSRs are less likely to raise competition concerns. However, Chen and Shaffer (2014) show that, under specific conditions, this is not necessarily the case. In particular, a seller with market power may find it profitable to offer MSRs, but not fully exclusive dealing contracts, when its intent is to nakedly exclude.

Their analysis builds on Rasmusen et al. (1991) and Segal and Whinston (2000a), sharing the same assumptions concerning the existence of scale economies, of multiple buyers and of a first-mover advantage. However, there are three key differences. First, they restrict contractual offers

¹⁶ For an extensive analysis, see Ide et al. (2016).

¹⁷ In Chapter 2 (and in particular in Sections 2.3.2 and 2.3.3), we discuss market-share discounts, which differ from the contracts discussed here in that they do not involve a commitment by the buyer, but consist of an offer by the seller to award a discount if *ex post* the volumes purchased during a certain period satisfy a certain condition, for instance they amount to a certain percentage of the requirements of the buyer.

to be uniform and allow buyers to coordinate their decisions. Hence, coordination failures cannot generate exclusion. Second, they assume that the fixed entry cost is uncertain when the contract is offered and buyers make their decisions. (The fixed entry cost is distributed between zero and an upper bound that guarantees that entry occurs for sure when no contract is offered by the incumbent.) Third, the contract involves not only a lump-sum compensation but also a commitment on the price at which the transaction with the incumbent will take place. The paper shows that, in this environment, an additional externality arises that can be exploited by the incumbent through MSRs; such externality does not appear when the contract is fully exclusive, as we discuss next.

When the contract is fully exclusive, the payoff of a buyer that signs *does not* depend on whether entry occurs. This is because that buyer will pay the contractual price on its entire requirement. Note that, in a setting in which coordination failures are assumed away, buyers can always jointly reject, induce entry for sure, and enjoy the surplus associated with the post-entry competitive price. Then, in order to sustain an equilibrium in which all the buyers accept the contract, the incumbent should 'break' joint rejections, by offering a compensation large enough to induce buyers to unilaterally deviate. If an incumbent is restricted to only making homogeneous offers, the incumbent should offer an amount that compensates each buyer (at least) for the loss suffered when paying the contractual price rather than the post-entry competitive price. However, as we have already discussed in Section 3.2.1, the profit obtained from each buyer, if she accepts the contract, is lower than the minimum compensation that a buyer requires (due to the deadweight loss arising for any contractual price that exceeds the incumbent's marginal cost), and offering fully exclusive contracts turns out to be unprofitable for the incumbent.

With MSRs, by contrast, even if a buyer accepts the contract, some units of purchase are free and can be supplied by the entrant at the competitive price, if entry occurs. Hence, in the case of MSRs, the payoff of a buyer that signs *does* depend on whether entry occurs. In turn, the probability of entry depends on the number of buyers that accept the contract. When that number is higher, the demand available for the entrant is more limited and the probability that entry costs are low enough to make entry profitable is smaller. Then, with MSRs *the payoff of a buyer that accepts the contract depends on the decision of the other buyers*. Recall that, absent coordination failures, to sustain an equilibrium in which all the buyers accept the

contract the incumbent must 'break' joint rejections. With MSRs, the incumbent must compensate the loss that a buyer suffers when, instead of paying the competitive price, it unilaterally accepts the contract and pays the contractual price on the units involved in exclusivity, while paying the competitive price on the free units when entry occurs (with a probability determined by a single buyer having accepted the contract). However, when everybody accepts exclusivity, buyers suffer an additional loss that is not compensated. This is because the decision of all the other buyers to accept exclusivity further decreases the probability that entry occurs and generates a new externality that is exploited by the incumbent. Indeed, the reduction in the probability of entry increases the profits that the incumbent obtains from each buyer and makes it profitable to offer contracts with MSRs.

Note that both the assumptions of random entry cost and price commitment are key to generate such an externality. The random entry cost ensures that the decision of each buyer to accept the contract smoothly decreases the probability of entry. Under a deterministic (that is, certain) entry cost, instead, there is always a buyer whose decision to accept makes the probability of entry switch from one to zero. The compensation that the incumbent should offer (to everybody) would be determined by the loss suffered by this buyer when paying the contractual price on her entire requirement (because the probability of entry is zero when she accepts) rather than the competitive price (because the probability of entry is one when she rejects). The decision of additional buyers to accept would not affect the probability of entry, and there would be no externality that can make the deal profitable for the incumbent. In a similar vein, there would be no externality without price commitment. Indeed if the price is determined *ex post*, that is, after entry takes place and the buyer pays the competitive price for the free units, then the incumbent would increase the price for the captive units to the level such that the average price over the entire buyer's requirement is equal to the monopoly price. Then, without price commitment, the buyer would pay the monopoly price irrespective of the probability of entry and of the decision of the other buyers. Hence, the externality set out above would no longer materialise if the incumbent was unable to commit to its prices *ex ante*.

Finally, Chen and Shaffer's model predicts that even though entry is not fully deterred, the use of MSRs contracts is welfare-detrimental: the probability of entry is lower than it would be absent any contract, buyers pay more and some output is produced by a less efficient producer (the incumbent).

Buyer power and asymmetric buyers The earlier analysis considers equal-sized buyers whose individual demand is insufficient to ‘sponsor’ entry.

If market demand was less fragmented, so that *each* buyer was sufficient to enable profitable entry, none of the exclusionary results discussed in Section 3.2.3.1 would apply and exclusive contracts could not be used to exclude a more efficient rival. In this respect, the existence of buyer power limits the incumbent’s ability to exclude by using exclusive contracts.

Let us consider next the existence of *asymmetric* buyers. The insights discussed so far extend to this environment as long as no buyer alone is sufficient to attract entry. When, instead, the individual demand of the large buyer(s) enables entry, it is still possible for the incumbent to exclude a more efficient rival. However, the incumbent must *fully* compensate the large buyers, or more generally buyers whose contribution to the entrant’s success is particularly important. Then, exclusion becomes more costly, in some cases unprofitable, as compared to the case of symmetric buyers (see Section 3.4.3.2 for the analytical treatment).

To see the intuition consider again the two-buyer setting. The large buyer – anticipating that its individual demand triggers entry – will reject exclusivity even when she expects the other buyer to accept, unless she is *fully* compensated. The main implication for the results is that the incumbent cannot take advantage of coordination failures to exclude in a ‘cheap’ way. Hence, if contractual offers must be homogeneous, ‘exclusionary equilibria’ do not exist: the full compensation for the large buyer should be offered to everybody and this clearly makes exclusion unprofitable for the incumbent.

If discrimination is possible, the incumbent can target the full compensation to the large buyer only, offering no compensation to the small one. Provided that the monopoly profits extracted from both buyers exceed the full compensation to the large one, offering such contracts is profitable for the incumbent, and exclusion arises at the equilibrium. However, when it exists, the ‘exclusionary equilibrium’ is unique, differently from the case of symmetric buyers where exclusionary equilibria were characterised by different levels of the compensation, including very modest ones. Also, the compensation that the incumbent must offer is larger than the one characterising the less favourable ‘exclusionary equilibrium’ in the symmetric case. In both respects, when buyers are asymmetric exclusion becomes more costly for the incumbent.

Similarly, when offers are sequential (for example, because buyers make their decisions at different times), the incumbent does not manage to

exclude at zero cost, but the unique ‘exclusionary equilibrium’ – when it exists – entails a full compensation paid to the large buyer.

Finally, note that in both cases (discriminatory offers and sequential offers) the condition for the existence of the ‘exclusionary equilibrium’ is that the monopoly profits obtained from both buyers exceed the full compensation to be paid to the large one. This condition is the less likely to be satisfied the larger the size of the big buyer. This is another channel through which asymmetry hinders exclusion.¹⁸

The above analysis assumes that all the bargaining power is held by the incumbent, who makes take-it-or-leave-it offers to buyers when it negotiates the exclusivity agreement. This assumption implies that the large buyer obtains the compensation that makes her *indifferent* between accepting exclusivity (and paying the monopoly price) and rejecting it (and paying the post-entry competitive price). However, a buyer that is large and whose individual demand suffices to make entry profitable is likely to have some bargaining power when negotiating with the incumbent. When this is the case, the large buyer receives a compensation that makes her *strictly* better off when she agrees on exclusivity with the incumbent than in the case in which the deal is not agreed upon and entry occurs. Effectively, the incumbent and the large buyer create a ‘coalition’ to extract rents from the remaining small buyer(s). By agreeing on exclusivity, the incumbent and the large buyer limit upstream competition which enables the incumbent to extract monopoly profits from the small buyer(s); it is precisely rents extraction from the small buyer(s) that makes the joint profits of the coalition larger under exclusivity than in the case in which no exclusivity is in place and entry occurs. When the incumbent holds all of the bargaining power, the large buyer obtains the part of the coalition profits that makes her indifferent between accepting and rejecting exclusivity. However, when the large buyer has some bargaining power, she appropriates a larger part of the coalition profit, the more so the stronger her bargaining power. In the limit, if the large buyer has all of the bargaining power, she appropriates the entire profits of the coalition. We develop in Section 3.4.3.3 the formal analysis of the case in which the large buyer has some bargaining power and we discuss in Section 3.5 the policy implications regarding this case.

¹⁸ In Johnson (2012) buyers’ asymmetry does not concern size. In an environment where consumers are uncertain about the quality of the entrant’s product, the incumbent may induce into exclusivity only the buyers/retailers that have enough reputational capital to legitimise the entrant. This does not deter entry entirely, but postpones the moment when consumers learn the quality of the entrant, thereby limiting its competitiveness in initial periods and allowing the incumbent to make larger profits.

Demand-side scale economies In the scenarios discussed so far, the critical factor leading to the conclusion that exclusive contracts can be used to exclude more efficient rivals is the presence of *scale economies*. Scale economies make the success of entry depend on the number of free buyers and generate the negative externality that one buyer's decision to purchase from the incumbent exerts on the others by making entry less profitable. The earlier analysis considers scale economies arising from the supply side, captured by the presence of fixed entry costs. A similar externality also arises if there are *demand-side scale economies*, for example, via *network externalities*: for a given population size, if a sufficiently large fraction of consumers choose the incumbent's product, the appeal of the rival's product (or network, or platform) will be limited, in turn prompting the remaining consumers to also choose the incumbent's network. This case has been studied by Doganoglu and Wright (2010). They consider competition between the incumbent and a more efficient rival, more efficient in the sense that the quality of the rival's network is superior to the incumbent's for equal number of users. The incumbent enjoys a first-mover advantage and can make an initial offer to a subset of users; in the following period, both the incumbent and the rival make their offers to the users still available. The main insight of the paper is that 'multi-homing' – that is, the possibility for consumers to buy simultaneously from more than one firm – is a key factor for exclusive dealing contracts to play a role in exploiting the above externality and lead to inefficient exclusion.

Consider first the case in which multi-homing is not allowed so that users either join the incumbent's network or the rival's. In that case, by offering a low price to a majority of users in the initial period (sufficiently low to make each of them indifferent between buying from the incumbent and buying from the rival at full size in the following period) the incumbent manages to build a customer base that is large enough to offset, through network externalities, the intrinsic advantage of the rival when competition for the remaining users takes place in the later period. Then, the rival is foreclosed, while the incumbent manages to charge a relatively high price to second-period users, which makes the strategy overall profitable. Note that a simple introductory price, without any commitment on the side of the buyer, suffices to exclude the rival.

Consider now the case in which multi-homing is feasible. If the incumbent offers simple introductory prices, without an exclusivity clause, exclusion cannot take place. The intuition is that users that accept the initial offer from the incumbent find it worthwhile to multi-home and buy also from the rival in the second period, so as to enjoy the benefit of

its superior network, while the remaining users buy exclusively from the rival. Hence, the possibility to multi-home cancels out any installed base of customers the incumbent can take advantage of in the second period and prevents the incumbent from extracting rents from second-period users. In turn, this makes it unprofitable for the incumbent to offer appealing introductory prices in the initial period. If, instead, the incumbent offers an introductory price with an exclusivity clause – that is, with the additional constraint that users that accept cannot buy also from the rival – then remaining users will choose to multi-home in the following period. This will allow them to reach exclusive users through the incumbent’s network and at the same time to benefit from the superior quality of the rival’s network. Since there is no competition for multi-home users, prices will be sufficiently high in the second period, which allows the incumbent to extract some rents from multi-home users in the second period. This, in turn, allows the incumbent to offer introductory prices (with exclusivity) that induce enough users to accept exclusivity in the first period. Note that, under multi-homing and exclusive introductory offers, the more efficient entrant is only *partially* foreclosed, as it cannot sell to users that accepted the incumbent’s introductory offer. Nevertheless, the outcome is inefficient, because the rival’s superior quality is not fully realised.

3.2.3.2 The role of downstream competition

The basic framework we have presented so far has interpreted buyers as final consumers. However, exclusive contracts are typically signed with firms (see the cases reviewed in Section 3.6 and Section 3.7). In this section, we assume that buyers are downstream firms (or retailers) that use the input bought either from the incumbent or from the potential entrant to turn it into a final output for end-consumers. The intensity of downstream competition is a crucial determinant of the scope for anti-competitive exclusive dealing. In particular, the main insight of this section is that, where scale economies are important and intermediate buyers are fragmented, exclusive contracts written between an incumbent and intermediate buyers are *more likely to be a concern if downstream competition is sufficiently weak*. By contrast, when downstream competition is fierce, different economic models lead to different predictions as to whether exclusion should be easier or not, as we shall discuss below. (See Section 3.4.3.4 for a formal analysis of the general framework presented here.)

To see the intuition behind the main insight just alluded to, consider first the case where buyers are downstream firms that operate in independent (separate) markets, so that downstream competition is absent. In this case the result is the same as the one we discussed within the basic framework with buyers interpreted as final consumers: if scale economies are strong enough and the input demand of a single buyer is insufficient to attract entry, the incumbent can exploit externalities among buyers to profitably exclude through exclusive contracts.

Suppose instead that downstream firms are identical and are Bertrand competitors (that is, they sell homogeneous products and compete in prices) and consider the case where one buyer has accepted exclusivity while the other has not. The more efficient entrant anticipates that, if it enters the market, it will be able to supply the free buyer at a lower price than the one the incumbent offered to the buyer that signed up for the exclusive contract. In this environment, where downstream competition is quite fierce, using a cheaper input than the rival provides the free buyer with a very strong competitive advantage: it will have a lower marginal cost and will be able to capture the entire downstream market. Thus, input demand from the free buyer alone will be sufficient for the entrant to cover its fixed entry costs. In other words, entry 'sponsored' by a single buyer is profitable in this setting.¹⁹

We have just considered the limit benchmark of Bertrand competition with homogeneous products. However, the basic intuition extends more generally to imperfect competition. The key message is that when downstream buyers compete, using a cheaper input than the rival allows to steal part of the rival's business. The more intense downstream competition, the stronger this business-stealing effect, the larger the input demand generated by the free buyer when it competes with a rival that has committed to buy from the incumbent. If downstream competition is sufficiently intense, the demand of the free buyer is large enough to cover the fixed costs of entry, thus making entry a profitable strategy. Hence, the fact that one buyer accepts exclusivity does not prevent the other from dealing with the entrant: the externality that was the core of the earlier analysis vanishes.²⁰

This discussion emphasises a first channel through which downstream competition affects the likelihood of profitable exclusion. More intense

¹⁹ This discussion refers to the case where upstream firms use linear wholesale prices. When upstream firms use two-part tariffs the analysis becomes more subtle, as two-part tariffs allow for a richer array of deviating strategies. See Wright (2009).

²⁰ This is the main insight in Fumagalli and Motta (2006).

competition decreases the number of buyers that the entrant needs to supply to be viable and, consequently, increases the number of buyers that the incumbent must compensate to exclude. Indeed, when downstream competition is sufficiently intense, *every* buyer is pivotal for entry and the incumbent must compensate *every* buyer to exclude. This effect decreases the likelihood of profitable exclusion through exclusive contracts, as exclusion would be too costly.

However, there is also a second channel through which downstream competition affects the likelihood of exclusion. Downstream competition has an impact on the profit that a buyer earns upon rejecting exclusivity (that is, when it deals with the entrant); this in turn affects the extent to which the incumbent needs to compensate a buyer in order to sign up to exclusivity. In this respect the role of downstream competition is not clear-cut.

On the one hand, fierce downstream competition may pass through much of the benefit of lower input prices, thereby eroding the profits of the free buyer and making it cheap for the incumbent to induce all buyers to sign up. Think for instance to the case where downstream buyers are Bertrand competitors, which we have just discussed: if one buyer rejects exclusivity while the other accepts, the former will pay a (slightly) lower price for the input. The free buyer captures the entire downstream market, thereby triggering entry, but its profit is close to zero (since the very small cost advantage will feed through into very small downstream margins, given that both competitors will price at cost). As a result, the incumbent can exclude by offering almost-zero compensation.^{21,22} In this case it is in the joint interest of the incumbent and the downstream buyers to agree upon exclusivity. By preventing entry, this limits upstream competition and keeps final prices high, thereby allowing more rents to be extracted from final consumers and increasing the surplus available to the firms for signing the contract.

²¹ See Simpson and Wicklegren (2007) and Abito and Wright (2008). In the same vein, intense downstream competition facilitates exclusion if the incumbent could resort to more sophisticated contracts that allow it to commit to sell cheaply to a party that signs up for exclusivity whenever this competes with a buyer that rejected exclusivity, thereby reducing the profits that the free buyer would expect to earn.

²² This effect of downstream competition, which facilitates exclusion, does not arise in the models developed in Chapters 1 and 2 because in that setting the incumbent does not enjoy a first-mover advantage and cannot lock-in buyers before the rival materialises and makes counter-offers. Then tough downstream competition, by making a single buyer sufficient for the rival to achieve efficient scale, unambiguously makes it less likely that predatory pricing and rebates can lead to exclusion.

On the other hand, one can also identify situations where the (sole) free buyer may appropriate (some of) the benefits of using a cheaper input, and these benefits are enhanced by more intense downstream competition. First, think back to the case where downstream buyers are Bertrand competitors but there are (small) fixed costs necessary to be active in the downstream market. In this case using a cheaper input forces a party that signs up for exclusivity out of the downstream market and allows the free buyer to monopolise it. Exclusion would be unprofitable as, to block rejections, the incumbent should offer every buyer the (vertically integrated) monopoly profits.

Second, suppose that the efficiency gap between the entrant and the incumbent is wide. The entrant has an incentive to set a wholesale price to the free buyer that is well below the incumbent's price for the signer. This allows the free buyer to make a sufficiently large profit, and may make exclusion too costly for the incumbent. This example indicates that it is not necessary that competition with the free buyer forces the signer out of the market for downstream competition to hinder exclusion.

Further, suppose that the (sole) free buyer, being the entrant's unique access to final consumers, has some bargaining power in negotiating with the entrant.²³ Then the free buyer manages to extract some of the surplus created when entry takes place. If this bargaining power – and thus the payoff of the free buyer when rejecting exclusivity – is large enough, exclusion becomes too costly for the incumbent.²⁴

Finally, one would reach a similar conclusion in the case where downstream buyers compete under capacity constraints, thus allowing (price-cost) margins not to be completely squeezed and in turn allowing the free buyer to make a sufficiently large profit. This last situation suggests that the effect of downstream competition on the likelihood of exclusion may exhibit non-monotonicities. In simpler words, increasing the intensity of downstream competition makes it harder to exclude when downstream competition is sufficiently weak to start with, as the dominant effect is the first channel we highlighted. But when downstream competition is already sufficiently strong, fiercer competition

²³ The papers mentioned above assume instead that upstream suppliers can always make take-it-or-leave-it offers to buyers/retailers.

²⁴ See Johnson (2012). He also notes that, for a given (positive) bargaining strength, exclusion is unprofitable for the incumbent if there are sufficiently many downstream retailers that the incumbent should compensate to sustain exclusion. This may be another way to interpret intense downstream competition.

may instead facilitate exclusion, because the pass-through effect may prevail.

The final conclusion that one can draw from this analysis is that, when downstream competition is strong, the picture is complex. Whether the incumbent can use exclusive contracts to exclude depends very much on the specific industry characteristics. One robust conclusion is that the basic ‘naked exclusion’ mechanism is more likely to lead to inefficient exclusion when downstream competition is weak. As for the application of these models to real world cases, as we note in Section 3.5, assessing the degree of competition is not a simple task in many industries. While there are some features which are typically associated with an industry or a market being more (or less) competitive, a detailed competitive assessment is likely to be required, focusing on any specific circumstances.

3.2.4 When Exclusivity can be Breached upon or Renegotiated

So far we have (implicitly) assumed that exclusivity is always binding without exceptions. However, this assumption is not realistic. When two parties write a contract, they may include in the contract itself what consequences would arise should one of the parties decide to terminate the contract or not fulfill its obligations. Therefore, we investigate in Section 3.2.4.1 what theory predicts if the exclusive contract already includes a clause specifying contractual damages to be paid to the incumbent in case the buyer wanted to breach exclusivity and purchase from the entrant instead. We then analyse in Section 3.2.4.2 the case where there is no *ex ante* provision for breach of contract, but rather the exclusive clause can be renegotiated *ex post* – for instance because after entry has materialised it is in the joint interest of all the parties (incumbent, entrant, buyer) not to enforce the clause and let the buyer buy from the entrant rather than from the incumbent.

3.2.4.1 Allowing for penalty fees for exclusivity breach

In this section we allow exclusive contracts to include *contractual damages*, that is, penalty fees (already established in the initial contract) that the buyer would pay to the incumbent if she breaches exclusivity and buys from an alternative supplier.²⁵ The discussion that follows will show that this type of contract can be used by the incumbent to benefit from the presence of

²⁵ Purely exclusive contracts can be interpreted as the limit case where contractual damages are infinitely high.

a more efficient rival by accommodating entry and *extracting through the penalty fee the efficiency gains that the entrant brings into the market*. From an economic perspective, the contract thus implements a transfer from a party outside the contract (the entrant) to the contracting parties. (The following discussion is mostly based on Aghion and Bolton (1987). See also Section 3.4.4 for an analytical illustration.)

To see the intuition, suppose there is a single buyer. Imagine also that such a buyer has accepted an exclusive contract, which also sets a penalty fee payable if exclusivity is breached. If the entrant enters the market, the buyer will prefer breaching the exclusivity clause if the entrant offers a sufficiently low price: the buyer can be better off by sourcing the product from the entrant at a low price, and pay the penalty fee to the incumbent, than by sticking to the incumbent's higher price. Of course, the higher the penalty, the lower the entrant's price that is required to induce a breach of exclusivity. Now, suppose the incumbent offers an exclusive contract with a penalty fee such that the price that induces exclusivity breach is just above the price at which the entrant breaks even. Say the buyer accepts the exclusive contract.²⁶ In spite of this contract being signed, the entrant decides to enter the market. It anticipates that the buyer will breach exclusivity; hence the entrant will sell to the buyer and will make slightly positive profits. In turn, it is profitable for the incumbent to offer this contract as it cashes in the penalty fee and it appropriates the surplus generated by the entry of a more efficient rival (so-called 'efficiency rent extraction'). Note that in this setting with a single buyer, a purely exclusive contract would *not* be signed (the Chicago School critique discussed earlier would apply). Hence, using exclusive contracts with penalty fees and engaging in rent extraction is the optimal choice for the incumbent.

On the other hand, if we replaced the single buyer with multiple buyers, and in the presence of scale economies, it is not obvious which type of contract is more profitable for the incumbent. The relative preference of one over the other depends on (i) how large the entrant's efficiency advantage is (and thus how large the efficiency rents that the incumbent can extract through a penalty fee are); and (ii) how significant scale economies are (that is, how profitable it is for the incumbent to exclude the more efficient rival through a purely exclusive contract).²⁷

²⁶ It is easy to show that the contract can be precisely designed in such a way that the buyer's payoff if she rejects is equal to (or slightly lower than) the buyer's payoff if she accepts and breaches exclusivity. See Section 3.4.4 for a formal treatment.

²⁷ For a deeper analysis see Segal and Whinston (2000a).

The above discussion refers to an environment where the contracting parties know for sure the characteristics of the future entrant, already at the time of signing the contract. In particular, the entrant's cost structure is known to all, so that the incumbent can establish a penalty fee at the exact level that allows efficiency rents to be extracted fully (that is, at the level whereby entry follows, the buyer breaches exclusivity and the entrant makes an infinitesimally small – but positive – profit). In this setting exclusive dealing contracts redistribute surplus in favour of the contracting parties; they are efficient for society as they allow the maximum total welfare to be achieved, namely the production by the most efficient supplier (regardless of how the rents are distributed).

Inefficiencies instead arise if, at the contracting stage, parties only know the probability distribution of the entrant's cost structure (that is, parties are unsure about the exact cost level of the entrant, they only know that it can take different values and they know the probability that each of these possible cost levels will realise). Hence, at the time of writing a contract, the incumbent cannot set a penalty fee that extracts the full efficiency rent, targeted to the specific entrant's characteristics. It has to set a penalty fee that depends on the *expected* cost of the entrant. It turns out that the penalty chosen by the incumbent makes the probability of entry inefficiently low: for certain cost values of the entrant, entry would have been efficient but does not take place because the penalty fee is too high. Note that such inefficient exclusion is a consequence of the uncertainty on the entrant's cost structure, rather than of a fully intentional incumbent's strategy (in fact, for such cost levels, the incumbent 'regrets' excluding the entrant *ex post*, as it is unable to extract an efficiency rent from it).²⁸

Jing and Winter (2014) extend the above setting by allowing the incumbent and the entrant to compete not only downstream, when they sell to the buyer, but also upstream, in the process of acquiring an essential input (assumed to be supplied by a producer with market power).²⁹ The authors show that in this environment an exclusive contract (between the incumbent and the buyer), which includes penalty fees upon exclusivity breach, may allow *rents to be extracted from the upstream supplier too*. Indeed, the higher the penalty fee, the lower the price that the entrant

²⁸ Regret takes place also when the realised cost is lower than the expected one and the chosen penalty fee is too low.

²⁹ Their analysis is inspired by a real-world case from Canada (*Nielsen*) that we will discuss in Section 3.7.6.

is forced to set to induce the buyer to breach and the lower the profits that the entrant makes in the downstream market. In turn, this implies that the entrant will be less aggressive when bidding for the upstream input, thereby leaving fewer rents to the upstream supplier. In a setting without uncertainty (that is, the entrant's cost is known to all) the optimal penalty fee will be chosen in such a way to (almost) cancel out the entrant's downstream profits. Then the entrant will only bid very low for the upstream input, thereby reducing (almost to zero) the profit of the upstream supplier. The exclusive contract therefore implements a rent transfer in favour of the 'buyer-incumbent coalition', but it does not produce inefficiencies. When, instead, there is uncertainty about the entrant's cost at the contracting stage, the optimal penalty fee will be such that inefficient exclusion may occur *ex post* consistently with our earlier discussion.

3.2.4.2 The role of contract renegotiation

In the previous sections we have explored the anti-competitive effect of exclusivity under the implicit assumption that the exclusive contract *cannot be renegotiated*. This means that, once the contract has been signed, the parties cannot change the initial agreement. Then, if enough buyers have signed up for exclusivity, the entrant will decide to stay out of the market because it anticipates that in case of entry it will not be able to attract enough demand to make its activity viable.

However, entry would be efficient from a societal perspective (since the entrant has a lower cost of production by assumption). This means that the joint surplus of the incumbent, the entrant and the buyers when entry takes place is larger than in the case where there is no entry. Therefore, if entry occurs, there is scope to renegotiate the initial agreement. The parties could find a mutually satisfactory outcome, whereby the buyers source their supply needs from the entrant; and the additional surplus generated by this efficient entry is distributed in such a way that each party is at least as well off as if the entrant had stayed out of the market.

Therefore, if amending the initial contract (that is, renegotiating) is costless (or entails sufficiently low costs), it is in the interest of all parties to do so. In such a case, however, the foreclosing effect of the exclusive contract vanishes: anticipating that the contract will be renegotiated, the entrant would not decide to stay out of the market even

though it observes that an exclusive contract has been signed by all the buyers.^{30,31}

This means that exclusive dealing may lead to entry deterrence only if renegotiation costs are sufficiently high. A possible source of renegotiation costs is the existence of transaction costs, as (re)negotiation activities may entail substantial effort, time and resources. They may also require legal advice, uncertainty and possibly (lengthy) judicial review or disputes. Moreover, delaying consumption and production until a new agreement is reached may be an additional source of costs. Renegotiation may be difficult also if the buyer customises its production process to the input supplied by the incumbent firm, thereby increasing the cost of switching to an alternative supplier.³²

Note that renegotiation may be particularly challenging under the presence of a *fragmented demand side*. First, the entrant needs to convince a large number of buyers to change the initial contract in order to achieve a viable scale.³³ Second, small buyers may have weaker incentives to renegotiate the initial agreement as compared to large buyers: they benefit less from efficient entry and may be unwilling to bear the renegotiation costs. This discussion highlights an additional reason why the existence of *buyer power* may limit the anti-competitive effect of exclusive dealing. Large buyers, even though none of them is large enough to attract entry, may be willing to bear renegotiation costs, thereby making it impossible for the incumbent to exclude through exclusive contracts. Asymmetries across buyers may also limit the scope for exclusion. Consider a market with two buyers, one larger than the other. Imagine that the large buyer is willing to bear the renegotiation costs, while the small buyer is not. The fact that

³⁰ A similar argument applies when efficient entry is deterred by too high contractual penalties à la Aghion and Bolton (1987) (see previous section). When renegotiation is costless, if entry occurred it would be in the mutual interest of the incumbent and the buyer to decrease the penalty and let breaching of exclusivity occur. Anticipating this, the entrant would decide to enter the market whenever it is efficient to do so.

³¹ Simpson and Wickelgren (2007) show that exclusive dealing may still be signed in equilibrium and be welfare-detrimental, even though costless renegotiation is feasible. However the negative effect on welfare is not due to entry deterrence but to allocative inefficiencies. This result arises when downstream competition is fierce. The idea is that fierce downstream competition, because of the pass-through effect, allows the incumbent to achieve exclusivity in spite of a very small compensation. Efficient entry takes place because one buyer will have the incentive to breach exclusivity and deal with the entrant. However, this buyer will set the monopoly price in the downstream market, thereby generating an allocative inefficiency.

³² See Milliou (2008) for examples of such technological lock-in.

³³ See Dewatripont and Maskin (1995) for an argument along these lines in the context of financial contracting.

the large buyer is willing to renegotiate makes the small buyer prone to reject exclusivity even if she expects the other buyer to accept exclusivity: indeed the small buyer is confident that entry will occur anyway because the large buyer will renegotiate the initial agreement and their combined demand will then be sufficient to make entry profitable. This prevents the incumbent from exploiting coordination failures and makes it more costly to use exclusive contracts to deter entry as compared to the case where renegotiation costs are prohibitively high.

3.2.5 Competing for Exclusivity

The models discussed so far give the incumbent a *first-mover advantage*: the incumbent can make a long-term contract offer to buyers before the rival can react and make a counter-offer. However, there are many real-world situations where none of the competing agents has the possibility to make an offer before the others. One may believe that removing the first-mover advantage hypothesis and allowing also the more efficient rival to take part in the initial negotiation makes inefficient exclusion impossible: whatever the offer the incumbent makes to buyers, the more efficient rival can always make a more favourable counter-offer and avoid exclusion. In fact, in technical Section 3.4.5 we formally show that the incumbent can exclude the entrant even when the latter can also bid for exclusivity. Next, we provide the intuition underpinning this result.

Consider a situation in which the incumbent and the rival offer exclusive dealing contracts to the first buyer, who can accept either exclusive offer or reject both. Then the entrant decides on entry, which involves the sinking of a fixed cost. Next, the suppliers offer exclusive dealing contracts to the second buyer. Finally, once contracts decisions have been taken, active suppliers set prices consistently with the contracts that have been signed and transactions take place.

In this setting, the reason why exclusion of the more efficient rival may take place stems from the same interaction between an incumbency advantage – for example, the fact that the incumbent has already sunk the entry cost when offers to the first buyer are made while the entrant has not – and significant scale economies – which manifest themselves through the entrant being able to operate profitably only if it manages to supply both buyers. This is the same mechanism that is central to the models developed in Chapters 1 and 2 to rationalise predation and anti-competitive rebates. Both firms anticipate that the seller that signs up the first buyer into exclusivity will eventually sell to the second. Therefore, each firm's bid for

the first buyer depends on what the expected gain from serving the second buyer is. While the incumbent extracts the monopoly profits if it serves the second buyer (because the entrant will not enter if it does not get the first buyer), the entrant cannot get more than the duopoly profits (because the incumbent will compete for the second buyer also in case it did not get the first buyer). Only when the difference between the marginal cost of the incumbent and that of the entrant is sufficiently large and the entry cost is not too sizeable, will the entrant be able to secure exclusivity with the first buyer. Differently stated, exclusivity with the incumbent, by excluding the entrant, softens second-period competition and allows the incumbent to extract larger rents than the rival from the second buyer, that is, from an agent not involved in the initial negotiation. Then, exclusivity with the incumbent maximises the joint payoff of the agents contracting in the first period – that is, the incumbent, the entrant and the first buyer – thereby allowing the incumbent to make an offer to the first buyer that the rival cannot outbid.

There are some features concerning this result that need to be highlighted.

Profit sacrifice is needed when competing for exclusivity It is important to note for our discussion that, when exclusion takes place in this framework without a first-mover advantage, the incumbent *does have to sacrifice profits*. Indeed, in the models we consider, the compensation which is offered to the first buyer is such that the incumbent *makes a loss on it*.

The argument is the following. Recall that the entrant is more efficient than the incumbent over total production and it covers the entry costs when it supplies both buyers at a price equal to the incumbent's marginal cost. Imagine that the incumbent makes an offer to the first buyer that does *not* involve a loss. Then the entrant could profitably undercut that offer. By so doing, it would secure the first buyer and, later, also the second. Since the initial offer to the first buyer did not involve losses for the incumbent, by undercutting that offer the entrant would earn total profits that are larger than the ones obtained when it supplies both buyers at a price equal to the incumbent's marginal cost, thereby covering the entry costs. Hence, in order to discourage entry, the incumbent must suffer losses on the first buyer so as to limit so much the profits that the entrant makes on the first buyer – if it decides to undercut – that the total profits that the entrant expects to make do not cover the entry cost. Despite the loss suffered on the first buyer, the offer is overall profitable for the incumbent because, as we explained above, when the first buyer agrees on exclusivity, then entry

does not take place, and the incumbent will not suffer competition from the rival in supplying the second buyer. This allows the incumbent to extract larger rents than the rival from the second buyer (namely the monopoly rents rather than the post-entry competitive profits) and to cover the initial loss.

Exclusive dealing contracts versus exclusivity discounts In the model outlined in this section firms agree on exclusivity first and, in later periods, all transactions take place. Such a timing is meant to capture the long-term nature of exclusive dealing contracts which may typically cover several years. In this setting, the fact that the contract includes an *ex ante commitment on the side of the buyer* is critical for exclusion. The reason is the following. If the first buyer signs the exclusive contract and commits not to buy from alternative suppliers, then the entrant will desist from entering the market, anticipating that the profits earned on the second buyer are insufficient to cover the entry cost. Lack of entry allows the incumbent to make monopoly profits on the second buyer and to cover the initial loss. Absent the *ex ante* commitment on the side of the buyer, the entrant would instead enter the market: since buyers decide at the end, when all the offers are on the table, the more efficient entrant would always cover the entry costs by undercutting any offer that is overall profitable for the incumbent. Hence, the anti-competitive concern over exclusive dealing contracts – which also involve an *ex ante* commitment on the side of the buyer – is more severe than that over exclusivity discounts, all else equal.

Exclusive dealing and rent extraction from non-contracting agents We have discussed above that, in the model developed in this section, exclusion of the more efficient rival takes place because it softens second-period competition and allows the incumbent to extract larger revenues than the rival from the second buyer, that is, from the agent not involved in the initial negotiation. As a consequence, exclusion of the rival maximises the joint payoffs of the agents contracting in the first period and allows the incumbent to make an offer to the first buyer that the rival cannot match. A similar logic applies to other settings. Consider the discussion about the role of downstream competition in Section 3.2.3.2. When downstream competition is fierce and much of the benefit of lower input prices is passed through to final consumers – that is, to outsiders in the initial contracting – it is in the joint interest of the incumbent and the downstream buyers to exclude the (more efficient) rival. By limiting upstream competition, this allows final prices to be kept higher, thereby extracting more rents from final consumers and increasing industry profits. Since the payoff of

final consumers is not internalised by the contracting agents, the efficient outcome for the contracting agents is socially inefficient.³⁴

Another application of the same principles is provided by Jing and Winter (2014). In an extension of their basic model (discussed in Section 3.2.4.1) they analyse a setting where the incumbent and the rival compete to acquire a number of (non-rivalrous) inputs – on top of competing for downstream buyers – and no first-mover advantage is assumed in the upstream bidding. They show that competition for the inputs will result in exclusion of the rival, with all the upstream suppliers selling only to the incumbent, when (i) complementarity between the upstream inputs is sufficiently high, (ii) downstream competition is sufficiently intense and (iii) bids are restricted to be monetary units rather than contracts. The third condition rules out the possibility to use non-linear contracts sustaining downstream prices that maximise total industry profits. Under this restriction, the other two conditions ensure that exclusion is in the joint interest of the incumbent, the rival and the upstream suppliers. Strong complementarity implies that selling some inputs to the incumbent and others to the rival generates a large loss: all the inputs must be used together. Intense downstream competition implies that the benefits of selling all the inputs to both suppliers are mostly transferred to final consumers. Together these conditions make it in the joint interest of the contracting parties to allocate all the inputs to the incumbent.

3.2.6 Empirical Evidence

The empirical evidence on the effects of exclusive dealing contracts is quite scant, especially on the entry deterrence motive. Lack of detailed data and methodological difficulties may explain such a gap.

³⁴ The same mechanism arises in Asker and Bar-Isaac (2014), albeit within an infinitely repeated game. In their model, the incumbent manufacturer and the potential entrant supply several retailers, that in turn serve final consumers. When upstream entry occurs industry profits are reduced because the benefits of efficient entry are transferred to final consumers through lower retail prices. Then, the joint payoff of the incumbent supplier and of the retailers may be higher when the entrant is excluded and more rents are extracted from final consumers. This creates the incentive for the incumbent supplier and the retailers to sign contracts that exclude the entrant and transfer some of the (larger) industry profits to retailers. The focus on the paper is on contracts such as loyalty rebates or resale price maintenance that, differently from exclusive dealing contracts, do not involve an obligation from retailers to buy in exclusivity.

Marin and Sicotte (2003) use the event-study methodology to circumvent these problems and investigate whether anti-competitive concerns are justified. The authors analyse the ocean shipping industry in the 1950s, which was characterised by extensive litigation over the use of exclusive contracts.³⁵ The paper uses several events in this turbulent period, each affecting the probability that exclusive contracts with customers would be legalised, and studies the stock price reaction to such events. Note that according to the authors a shared view at the time was that exclusive contracts were used by shipping cartels to deter the entry by out-of-cartels firms. The empirical analysis finds evidence supporting this view. Indeed, some of the most important events caused a significant increase in shipping firms' stock returns, while decreasing the value of companies in net-exporting industries that relied on shipping extensively. The authors interpret these results as suggesting that the object of exclusive contracts in this industry was to increase the cartels' market power, not to enhance efficiency or to benefit customers.

A few recent papers rely on a structural econometric approach. Asker (2016) investigates the potential impact of exclusive dealing arrangements on the beer industry in Chicago, testing whether they have hindered access to efficient distribution channels to competing brewers.³⁶ The study uses a detailed description of the Greater Chicago distribution networks of 12 brewers, merged with data on the sales of beer from a major Chicago supermarket chain, including information about the price paid by the chain and by consumers for the units of beer sold. A further interesting industry feature was the existence of local regulations granting an exclusive territory to each distributor. This rich data set allows demand parameters to be estimated and then cost parameters for brewers and distributors. The implementation of the test exploits the fact that, in the Chicago beer market, all Anheuser Busch distributors are exclusive, half of the Miller ones are exclusive while all other brewers use shared distributors. The test thus compares the cost to brewers, other than Anheuser Busch and

³⁵ This paper explains that after years of litigation, the US Supreme Court in 1958 established that the use of such contracts by shipping cartels was illegal. A political struggle ensued, which was resolved in 1961 when a new legislation exempting these contracts from antitrust scrutiny was approved.

³⁶ The author notes that the beer industry is an interesting case to study because there is a limited number of large brewers and exclusive dealing arrangements are a common practice. In 1994, the year of analysis, Anheuser Busch produced 45 per cent of the beer shipped in the US, followed by Miller with 22 per cent. The Anheuser Busch Annual Reports for 1998 and 1999 reveal that in 1998 40 per cent of its beer was sold through exclusive distributors; by 1999 this had risen to 60 per cent.

Miller, of serving a market with a Miller exclusive to the cost of serving a market without a Miller exclusive, using the Anheuser Busch distributors as a control group to pick up market-specific effects. The rationale of the test is that the costs to brewers should be higher in the former case if exclusive dealing contracts foreclose access to more efficient distributors. This hypothesis is rejected by the data, indicating that foreclosure effects are unlikely to be present in the Chicago beer market.³⁷

The analysis of Nurski and Verboven (2016) on the car market in Belgium leads, instead, at least to some extent, to the opposite conclusion.³⁸ They collect a rich data set combining information on 1860 dealers (including address and brands of cars sold) with data on car sales by model within a given zip code, car characteristics and consumer characteristics (gender, zip code and demographics within each zip code). This allows various demand and cost parameters to be estimated. In particular, demand estimation reveals that consumers value dealer proximity when buying a car (mostly because they care about after-sale services). This implies that new entrants need a sufficiently granular distribution network in order to operate successfully in the market. According to the literature reviewed in this section, the fixed costs associated with the set up of a network, together with the strong scale economies that characterise car production, make the anti-competitive effect of exclusive contracts a possible concern in this environment. The counterfactual analysis performed in Nurski and

³⁷ There exist other papers seeking to estimate the effects of exclusive contracts in the US beer industry, with mixed results. For example, Chen (2014) estimated an entry model to study the effect of exclusive dealing between Anheuser Busch and its distributors on rival brewers' entry decisions and consumer surplus (including the degree of choice available to consumers) and found that any welfare improvement associated with such a ban would be very small.

Chen and Shieh (2016) exploited the introduction of a distribution agreement in 2007 that allowed the European InBev beer brands to join Anheuser Busch's exclusive-distribution network simultaneously across all geographic areas in the US. The paper found that InBev's market share increased by 6 per cent once its beers started to be carried by Anheuser Busch's exclusive distributors, without an increase in retail prices. The paper further found that '[t]he effect [on InBev's market share was] strongest for stores with a small sales area and that carry the largest number of brands, where product retail competition [was] fierce and a local distributor's promotional effort [was] more likely to make an impact'. By exploiting variation in Anheuser Busch's market share across stores, the paper rejected the hypothesis that InBev's increase in market share (after it joined Anheuser Busch's exclusive-distribution network) may have been driven by any Anheuser Busch's market power. The authors interpreted their results as evidence that products carried by exclusive distributors received greater promotional efforts through non-price channels.

³⁸ According to this paper, exclusive dealing in automobile distribution is prevalent in most European countries, averaging about 70 per cent.

Verboven (2016) partially confirms this view. The authors compute the notional equilibrium that would arise with a ban on exclusive dealing and find that such a ban would shift market shares from the larger European manufacturers to the smaller entrants. Consumers would gain from such a ban, but mainly because of the increased availability of different car brands within a shorter distance from home, rather than because of substantially lower prices.

3.3 Efficiency-enhancing Exclusive Dealing

Section 3.2 has illustrated the main insights of the literature investigating the potential anti-competitive effects of exclusive contracts. Another strand of the literature, that we will discuss in this section, has explored whether exclusive contracts are actually signed to realise efficiencies. Among these, one that has received a lot of attention is whether exclusive contracts may protect relation-specific investments against opportunistic behaviour by buyers that would lead to a so-called hold-up problem.³⁹ We begin by reviewing this argument, before turning to other potential efficiencies that arise through the use of exclusive contracts.

3.3.1 Exclusive Dealing and Protection of Relation-specific Investments

To see why exclusive dealing may matter for investments, consider a manufacturer that can invest in the relationship with its retailer in order to increase the quality of retailing services and the value of trade. For instance, it can provide (i) sales training for retail managers and employees; (ii) technical support for local advertising and promotion; (iii) information about potential customers; (iv) equipment for servicing and repair; (v) financing to build and furnish the retail outlet.⁴⁰ The

³⁹ The EFTA Surveillance Authority provided an informal definition of hold-up as follows: ‘A “hold-up” problem may arise when there is a risk that one party may act opportunistically once another party has made client-specific investments in for example equipment or training, or transferred substantial know-how to the other party. Free-rider and hold-up problems may be anticipated prior to the conclusion of the agreement and may result in underinvestment if they cannot be resolved by the use of vertical restraints.’ (*Posten Norge*, para. 689)

⁴⁰ These investments can be illustrated by various antitrust cases on exclusive dealing, as noted in Besanko and Perry (1993). Specifically, footnote 3 of this paper states: ‘In *Standard Oil Co. v US* (1949), Standard Oil of California provided financing for the construction of gasoline stations. In *FTC v Brown Shoe* (1966), Brown provided architectural, management, and accounting support to retailers. In *American Motor Inns*

benefits of such investments, however, are not typically confined to the investing manufacturer. If the retailer carries competing products, their manufacturers will also benefit from such investments. For instance, once technical and sales training provided by one manufacturer makes the retailer more effective at selling a given category of products, not only the investing manufacturer but also the producers of rival brands will benefit from it.

This example highlights that the investment may not just increase the value of trade between the retailer and the investing manufacturer (this effect is denoted by Segal and Whinston (2000b) as the *internal effect* of the investment) but it may also affect the value of trade between the retailer and alternative manufacturers (*external effect*). In particular, in the example we are discussing, the external effect is *detrimental* to the investing firm: by improving the quality of retailing services, the investment facilitates also the sales of rival brands, thereby making the investment less valuable.

This suggests why, in similar circumstances, exclusive dealing promotes the investment. Absent exclusivity the retailer carries several brands so that, when deciding how much to invest, the manufacturer takes into account both the internal and the external effect of the investment. Under exclusivity, instead, the retailer carries only the brand of the investing manufacturer so that there is no external effect to be concerned about. Since the external effect is detrimental to the investing manufacturer, ignoring it under exclusivity makes investment incentives stronger and induces the manufacturer to invest more as compared to the case where there is no exclusive dealing arrangement, with a likely positive impact on end-consumers and welfare. Note that the assumption that the investment is not contractible is crucial for this argument. If the manufacturer and the retailer could write an enforceable contract establishing the desired investment level and the associated reward for the investing agent, exclusive dealing would have no role in promoting investment.

The conclusion would be the opposite if the external effect of the investment deteriorates, instead of increasing, the value of the external relationship. Consider, for instance, a situation where the investment of

Inc. v Holiday Inn (1975), Holiday Inn provided a nationwide reservation and referral system for its franchisees. In addition, Holiday Inn provided training to its franchisees through its Holiday Inn University. In *FTC v Beltone Electronics Corp.* (1982), Beltone, a manufacturer of hearing aids, provided its retailers with local customer leads obtained through consumer responses to national advertising. Finally, in *Rohery Storage and Van Co. v Atlas Van Lines* (1986), Atlas provided driver and employee training for its franchisees.

an input supplier, which improves the compatibility between its product and the buyer's production process, makes it more costly for the buyer to use alternative inputs. In this case the external effect of the investment is *beneficial* to the investing supplier as it limits the competitive pressure exerted by rival input suppliers. Hence, ignoring the external effect when an exclusive contract is in place makes investment incentives weaker as compared to the case where there is no exclusivity, and the supplier is induced to choose a lower level of investment. In other words, in this case, exclusive dealing would stifle rather than promote relation-specific investments.

The above discussion suggests that the *sign of the external effect* (whether beneficial or detrimental to the investing agent) is an important determinant of the effect of exclusive dealing on investments. The *identity of the investing agent* is equally crucial. Indeed, in the examples proposed so far the investment is undertaken by the supplier. Let us consider now situations where it is the buyer (typically a retailer) who invests. For instance, the retailer may invest (or exert effort, devote time) in order to promote a particular brand at the expense of competing brands. In such a situation, the investment increases the value of trade between the retailer and the manufacturer of that specific brand, while it decreases the value of trade between the retailer and the producers of rival brands. The external effect of the investment is therefore detrimental to the investing retailer. As a consequence, an exclusive dealing arrangement with a specific manufacturer, by making the retailer ignore the external effect of the investment, will increase the retailer's promotional investment (or the focus) for the exclusive brand. Applying a similar logic one would conclude that exclusive dealing limits the retailer's investment if the investment increases both the internal and the external value.

The general conclusion that can be drawn from this literature is, therefore, that when the supplier (manufacturer) invests, exclusive dealing is more likely to promote investment when the external effect increases the value of trade between the buyer and alternative suppliers, thereby being detrimental to the investing manufacturer. Instead, when the buyer (retailer) invests, exclusive dealing is more likely to promote investment when the external effect decreases the value of the external relationship and for this reason is detrimental to the investing buyer (retailer).⁴¹

⁴¹ This discussion has mostly drawn from Segal and Whinston (2000b), Besanko and Perry (1993), Marvel (1982), and Sass and Gisser (1989). On this issue, see also Spier and

The literature discussed so far focuses on the investment-promotion effect of exclusive dealing while abstracting from the risk of foreclosure, as it assumes that the initial contract is always renegotiable.⁴² At the other extreme, the literature illustrated in Section 3.2 focuses on the risk of foreclosure (by assuming that the exclusive contract is not renegotiable) while abstracting from investment promotion, as it does not model the possibility that the contracting parties engage in relation-specific investment. Fumagalli et al. (2012) analyse a model where both investment promotion and the risk of foreclosure can arise simultaneously. They show that the *interaction* between these two effects provides interesting insights on the welfare effects of exclusive contracts. In particular, a contract that forecloses a more efficient supplier may be signed *precisely* because it fosters investment.

To see the intuition, consider a situation where an incumbent supplier offers a single buyer an exclusive contract which cannot be renegotiated. Absent any effect on investment, the Chicago School critique applies and an exclusive contract that leads to foreclosure of a more efficient rival would not be signed in equilibrium: since the buyer's loss from exclusivity is larger than the incumbent's gain, the incumbent could never obtain the buyer's acceptance in a profitable way. Now, let us take into account the possibility that the exclusive contract may stimulate investment which is specific to the buyer–incumbent relationship. Investment promotion, by increasing the value of trade between the incumbent and the buyer, mitigates the buyer's loss due to exclusivity and expands the incumbent's gain. If this effect is sufficiently strong, the buyer and the incumbent have a private incentive to sign the exclusive contract. However, if investment promotion is too weak to make the incumbent more efficient than the rival supplier, the decision to agree on exclusivity is welfare-detrimental, as it forecloses an efficient rival.

There are two main messages from the last paper. First, it has identified a new situation where exclusive dealing might lead to inefficient foreclosure, thus expanding the set of situations where a possible theory of harm exists. In particular, it shows that even absent fragmentation of buyers, scale economies and the other circumstances discussed above possibly leading to

Whinston (1995), De Meza and Selvaggi (2007), Vasconcelos (2014), Groh and Spagnolo (2004) and Segal and Whinston (2007). Section 3.4.6 will address the issue in a formal way.

⁴² The approach used by Besanko and Perry (1993) is slightly different because foreclosure issues are avoided by assuming that manufacturers face a perfectly elastic supply of potential retail outlets.

exclusion, it is investment promotion that may make the risk of foreclosure a likely concern. Second, this analysis suggests that the balancing exercise of any anti- and pro-competitive effects of exclusive dealing might be even more complex than already acknowledged, since the very existence of investment-promotion effects – rather than being only welfare-beneficial – might actually make it easier for exclusive dealing to foreclose efficient entry.⁴³

3.3.2 Other Efficiency Justifications

Exclusive dealing contracts may generate efficiency gains also by solving incentive problems between sellers/manufacturers and buyers/retailers in an environment where their relationship is affected by asymmetric information.

Think for instance of a retailer who can take action to switch consumers among competing brands. These actions are not observable by upstream manufacturers (that is, a moral hazard problem arises). When competing manufacturers distribute their products through a common (risk-averse) retailer, they will devise contracts in order to provide incentives for such a common retailer to favour their product. These contracts, however, may result in too much risk being transferred to the retailer, thereby generating distortions. This problem is solved if the retailer distributes the product of a single manufacturer, that is, under exclusivity.⁴⁴

3.3.3 Empirical Evidence

Just as in the case of anti-competitive exclusive dealing, the empirical evidence concerning exclusive contracts and investment promotion is quite scant. Heide et al. (1998) conducted a survey of managers responsible for distribution decisions in 147 firms in the industrial machinery and electric equipment industries. They find evidence that managers are more likely to use exclusive dealing when there is a potential that the services they provide to distributors also benefit rival manufacturers. Sass (2005) focuses on the US beer industry, by exploiting the 1996/97 Distributor Productivity-Brand

⁴³ Another reason why exclusive dealing, by fostering investment, may be welfare-detrimental has been identified by Spier and Whinston (1995) and Segal and Whinston (2000b). In a setting where renegotiation is costless and thus the risk of foreclosure is not a concern, these papers show that exclusive dealing can be used strategically by the buyer-incumbent coalition to extract more rents from the more efficient supplier. Since investment is sub-optimally high, total welfare is reduced.

⁴⁴ On this role of exclusive dealing, see Bernheim and Whinston (1998).

Equity Survey, which contains detailed information (including on sales, costs and promotional activities) for individual beer wholesalers. The paper finds that the level of national advertising by brewers is not a significant determinant of distributor exclusivity. However, exclusive dealing is found to be positively correlated with both the price charged by brewers and the quantity of beer sold by distributors, suggesting that exclusion is unlikely to be the motivating factor behind the use of exclusive dealing, while being consistent with the investment-promotion hypothesis. Asker (2016) and Chen and Shieh (2016), already mentioned in Section 3.2.6, carried out empirical research in the US beer industry and found that exclusive contracts may drive efficiencies (mainly through increased promotional effort).

3.4 Economic Models of Exclusive Dealing*

In this section we provide a formal treatment of different models of exclusive dealing. We begin (Section 3.4.1) with the formalisation of the Chicago School critique to the notion of anti-competitive exclusive dealing. Next, we deal with models that: (i) discuss the role of upstream competition (Section 3.4.2); (ii) assume the existence of multiple buyers and of scale economies (Section 3.4.3); (iii) discuss the role of contractual penalty fees upon exclusivity breach (Section 3.4.4); (iv) consider the case where the incumbent does not have a first-mover advantage but competes with the entrant for exclusivity of the buyers (Section 3.4.5); and (v) formalise in a simple way pro-competitive exclusive dealing via investment promotion (Section 3.4.6).

3.4.1 The Chicago School Critique*

To formalise the Chicago School's argument against the existence of any anti-competitive effects from exclusive dealing, consider the following simple game. In the first stage, the incumbent (denoted as firm I) offers the buyer a compensation $x \geq 0$ in order for her to accept an exclusive contract. In the second stage, the buyer accepts or rejects the offer. In the third stage, the potential entrant, denoted as firm E – after having observed whether a contract has been signed or not – decides whether to

enter (and sinks the entry cost f if it enters). In the last stage, active firms choose prices and sell (the same homogeneous good).^{45,46}

The entrant exhibits a lower marginal cost than the incumbent: $0 = c_E < c_I$. Demand is given by $D(p) = 1 - p$, with $c_I < 1/2$ (this condition restricts the cases to be considered in the price game, see below). Finally, $f < c_I(1 - c_I)$, which implies that, absent exclusive contracts, entry is profitable and welfare-beneficial.

We look for the sub-game perfect Nash equilibrium of this game, and solve the model backwards.

At the last stage of the game, if no entry has occurred, the incumbent is the only seller. It is easy to check that the optimal price, the corresponding profit and consumer surplus are:

$$p_I^m = \frac{1 + c_I}{2}; \quad \pi_I^m = \frac{(1 - c_I)^2}{4}; \quad CS(p_I^m) = \frac{(1 - c_I)^2}{8}. \quad (3.1)$$

The outcome is the same if entry occurred but the buyer signed the exclusive contract. The entrant makes zero sales in this case.

If instead entry has occurred and the buyer is free from any exclusivity obligation, then Bertrand competition implies that the market is served by the more efficient entrant at the price which equals the marginal cost of the incumbent: $p^e = c_I$.⁴⁷ The incumbent's payoff is zero while the consumer surplus is easily computed as $CS(c_I) = (1 - c_I)^2/2$.

At the previous stage, if the buyer is 'free', the entrant anticipates that future earnings are sufficient to cover the entry cost (that is, $c_I(1 - c_I) - f > 0$) and decides to enter the market. Instead, if the buyer signed the exclusive contract, the entrant anticipates that it will not sell any unit

⁴⁵ Note that implicitly we are making some assumptions which we shall discuss later on in the chapter. First, the exclusive contract cannot be renegotiated. Second, it does not include penalty fees upon breaching the contract. In other words, if the buyer has signed the contract and entry takes place, the buyer cannot buy any unit from the entrant. Third, the existence of a more efficient producer considering entry is common knowledge at the contracting stage.

⁴⁶ The Chicago School argument is based on the assumption that if entry occurs, there is *perfect competition* between firms I and E . Since it leads to the same implications and we want to cast the arguments in a game-theoretic way, we assume *Bertrand competition* instead.

⁴⁷ If the entrant was much more efficient than the incumbent it could be that the entrant's monopoly price, $1/2$, is lower than c_I . The assumption made above guarantees that this is not the case, and simplifies the analysis (assuming otherwise would not change the qualitative results, though).

and entry costs would remain uncovered. Hence, it decides to stay out of the market ($-f < 0$).

Next, we have to check whether the buyer accepts the exclusivity offer from the incumbent. She will do so if the compensation offered by the incumbent will offset the loss suffered paying the monopoly price instead of the price c_I prevailing under entry: $CS(p_I^m) + x \geq CS(c_I)$. In other words the buyer accepts if $x \geq CS(c_I) - CS(p_I^m) = 3(1 - c_I)^2 / 8 \equiv x_{\min}^b$.

Finally, let us analyse the incumbent's decision. By having the contract signed the incumbent obtains a payoff $\pi_I^m - x$, whereas without exclusivity it will have zero profits. Then, the maximum compensation that the incumbent is willing to pay is

$$x_{\max}^b = \pi_I^m = (1 - c_I)^2 / 4 < 3(1 - c_I)^2 / 8 = CS(c_I) - CS(p_I^m) = x_{\min}^b. \quad (3.2)$$

Therefore, it is not profitable for the incumbent to induce the buyer to accept the exclusive contract; entry will not be deterred.⁴⁸

3.4.2 Upstream Competition*

Consider the same game as in the previous section, but with one change: if entry occurs, firms compete in quantities rather than in prices, a weaker form of competition than price competition. Let us solve the game backwards.

Either if entry did not occur or if entry occurred and the buyer signed the exclusive contract, the analysis of the last stage of the game is unchanged.

If instead entry occurred but the buyer did not sign an exclusive contract, firms would compete à la Cournot. Each firm's problem is $\max_{q_i} \pi_i = (1 - q_i - q_j - c_I)q_i$, with $i, j = I, E$ and $i \neq j$.⁴⁹ Thus, equilibrium quantities are:

$$q_E^e = \frac{1 + c_I}{3}; q_I^e = \frac{1 - 2c_I}{3}. \quad (3.3)$$

⁴⁸ Note that the result that $\pi_I^m < CS(c_I) - CS(p_I^m)$ has a general validity, and does not depend on the specific demand function adopted in this example.

⁴⁹ The assumption $0 < c_I < 1/2$ ensures that competition results in both firms selling positive quantities.

Post-entry equilibrium prices, incumbent's profits and consumer surplus can be found by substitution:

$$p^e = \frac{1 + c_I}{3} > c_I; \quad \pi_I^e = \frac{(1 - 2c_I)^2}{9} > 0; \quad CS(p^e) = \frac{(2 - c_I)^2}{18}. \quad (3.4)$$

For future reference, it is also convenient to compute the entrant's profits and total welfare:

$$\pi_E^e = \frac{(1 + c_I)^2}{9}; \quad W^c = \frac{8(1 - c_I) + 11c_I^2}{18}. \quad (3.5)$$

As in the previous example, at stage 3 entry occurs if (and only if) the buyer is free.

At stage 2, the buyer decides on exclusivity. The minimum compensation required to sign the exclusive contract must cover the surplus lost paying the monopoly price instead of the post-entry price p^e :

$$x_{\min}^c \geq CS(p^e) - CS(p_I^m) = \frac{7 + 2c_I - 5c_I^2}{72}. \quad (3.6)$$

Note that under Cournot competition the post-entry price is higher than the incumbent's marginal cost c_I . This reduces the minimum compensation required by the buyer with respect to the case of Bertrand competition ($x_{\min}^c < x_{\min}^b$).

In turn, the maximum compensation that the incumbent is willing to offer amounts to the additional profits earned when entry is deterred:

$$x_{\max}^c \geq \pi_I^m - \pi_I^e = \frac{5 - 2c_I - 7c_I^2}{36}. \quad (3.7)$$

The incumbent's maximum compensation is reduced by Cournot competition ($x_{\max}^c < x_{\max}^b$). Since competition is weaker, the incumbent earns positive post-entry profits. Instead, in the Bertrand case, competition is so fierce that the incumbent's post-entry profits are driven down to zero. This makes the gain from having the contract signed (and deterring entry) larger.

It is profitable for the incumbent to drive the buyer's acceptance of exclusivity if $x_{\max}^c \geq x_{\min}^c$, or:

$$1 - 2c_I - 3c_I^2 \geq 0 \quad (3.8)$$

$$= (1 + c_I)(1 - 3c_I) \geq 0. \quad (3.9)$$

This condition is satisfied if (and only if) $c_I \leq 1/3$. In other words, under Cournot competition (that is, when competition is relatively

weaker than under Bertrand competition), the incumbent can profitably use exclusive dealing to deter efficient entry provided that the efficiency gap with the entrant is not too large.

3.4.3 Multiple Buyers and Scale Economies*

In this section, we modify the assumption that there is only one buyer: we assume throughout that there are several buyers and that the entrant needs to win a sufficient number of orders to reach minimum efficient scale. We examine the main model in Section 3.4.3.1, and then deal with three extensions: Section 3.4.3.2 considers the case of asymmetric buyers; Section 3.4.3.3 considers the case where there is one large buyer with bargaining power; and Section 3.4.3.4 the case where buyers are not final consumers but downstream firms which compete among themselves in the downstream market.

3.4.3.1 Symmetric buyers*

Let us modify the setting studied above as follows. In the first stage the incumbent offers exclusive contracts to two identical buyers, B_1 and B_2 whose individual demand is given by $D(p) = 1 - p$. In the second stage buyers non-cooperatively decide whether to accept or reject. The rest of the game remains unchanged.

The entry cost f is assumed to be too large for entry to be profitable if the entrant supplies only one buyer, but small enough for entry to be profitable if firm E supplies both customers:

$$c_I D(c_I) < f < 2c_I D(c_I). \quad (3.10)$$

Finally, recall that the monopoly profits extracted from both buyers are insufficient to cover the loss suffered by both buyers when paying the monopoly price instead of the competitive price c_I :

$$2\pi^m(c_I) = \frac{(1 - c_I)^2}{2} < \frac{3(1 - c_I)^2}{4} = 2[CS(c_I) - CS(p^m(c_I))] = 2x_{\min}^b. \quad (3.11)$$

Simultaneous and non-discriminatory offers Let us start by analysing the case where the incumbent must offer the same contracts to all the buyers, that is, contracts committing to pay the same compensation x in exchange for exclusivity. Then, buyers simultaneously decide whether to accept or reject this offer.

There exist two types of equilibria of the whole game. There exist ‘*exclusion equilibria*’ where the incumbent offers $x \in [0, \pi^m]$, both buyers sign, the entrant does not enter, and the incumbent sets the monopoly price $p^m(c_I)$. There exist also ‘*entry equilibria*’ where the incumbent offers $x \in [0, x_{\min}^b]$, both buyers reject, entry follows and the entrant charges the competitive price c_I .

Let us solve the game backward. At the last stage of the game, price decisions are straightforward. If entry did not occur, the incumbent charges the monopoly price $p^m(c_I)$ to all the buyers, irrespective of whether they accepted the exclusive contract or not. If entry occurred, the entrant supplies the free buyer(s) at the price c_I , if any, while the incumbent charges the monopoly price to the signer(s).

At stage 3, the entrant observes how many buyers accepted the exclusive deal and decides on entry. By assumption (3.10), it decides to enter the market only if both buyers rejected the incumbent’s exclusive contract.

Anticipating this, at stage 2, for a given compensation x , buyers take the following decisions: if $x > x_{\min}^b$, there exists a unique equilibrium where both buyers sign the exclusive deal; if $x \leq x_{\min}^b$, there exist two equilibria, one such that both buyers sign, the other such that both buyers reject.

To see the intuition, consider a situation where both buyers sign the contract. Entry does not occur and each of them pays the monopoly price, enjoying a payoff $CS(p^m(c_I)) + x$. If a single buyer deviates and rejects the contract, entry does not follow either as by assumption (3.10) the deviant buyer’s demand alone does not attract entry. Hence, the deviant buyer ends up paying the monopoly price and enjoys a payoff $CS(p^m(c_I))$. Then, for any $x \geq 0$, unilateral deviations are unprofitable and both buyers signing the contract is an equilibrium.

Consider now a situation where both buyers reject the contract and entry occurs. They end up paying the competitive price c_I . A buyer who deviates unilaterally and signs the contract pays the monopoly price. If the compensation offered by the incumbent is (weakly) lower than the loss suffered by a buyer when she pays the monopoly price instead of the price c_I (that is, is weakly lower than x_{\min}^b) the deviation is unprofitable. The above situation is an equilibrium. If instead, the compensation offered is above x_{\min}^b , the deviation is profitable and the candidate equilibrium is broken.

Let us now move to stage 1, where the incumbent offers the contracts. The optimal incumbent's choice depends on the buyers' continuation strategies. If, following any offer where the compensation x is (weakly) below x_{\min}^b , continuation equilibria are such that both buyers reject the contract, the incumbent is indifferent between offering any $x \in [0, x_{\min}^b]$. In all these cases entry takes place and the incumbent makes zero profits. Since offers must be uniform, in order to induce both buyers to sign and deter entry the incumbent should offer *every* buyer slightly more than x_{\min}^b . By assumption (3.11) such a deviation is not profitable for the incumbent.

If, instead, following any offer where $x \geq 0$, continuation equilibria are such that both buyers accept the exclusive contract, the optimal choice of the incumbent is $x = 0$. Entry will not follow and the incumbent will obtain the monopoly profits from both buyers. Equilibria where the incumbent chooses $x > 0$ are sustained by having the continuation equilibria following any offer $x' < x$ be such that no buyer signs. Then, the incumbent has no incentive to decrease the compensation because it would lose both buyers. Of course, there are no 'exclusion equilibria' where the incumbent offers more than π^m to both buyers.

Simultaneous and discriminatory offers Suppose now that the incumbent can discriminate the compensation across buyers. In this case 'entry equilibria' *do not exist*. The crucial point is that, in the model presented in these sections, even if it is not profitable for the incumbent to offer $x = x_{\min}^b$ to both buyers, it is profitable to offer $x = x_{\min}^b$ to a *single* buyer:

$$2\pi^m(c_I) = \frac{(1 - c_I)^2}{2} > \frac{3(1 - c_I)^2}{8} = [CS(c_I) - CS(p^m(c_I))] = x_{\min}^b. \quad (3.12)$$

Then, imagine that both buyers reject the exclusive contract. Entry occurs and the incumbent makes zero profits. Instead, if the incumbent deviates and offers (slightly more than) x_{\min}^b to one buyer and (slightly more than) zero to the other, both buyers sign. The reason is that the dominant strategy for a buyer who is offered slightly more than x_{\min}^b is to sign the exclusive deal. Given that one buyer signs and that individual demand does not trigger entry, the other buyer cannot do better than signing. Hence, following such a deviation, entry is deterred

and the incumbent extracts the monopoly profits from both buyers. Since $2\pi^m(c_I) > x_{\min}^b$, the incumbent's deviation is profitable.

Instead, 'exclusion equilibria' do exist. In particular, $x_1 = x_2 = 0$, followed by both buyers signing the contract, is sustained as an equilibrium by having both buyers signing the contract following any offer where $x \geq 0$. The incumbent has no incentive to increase the compensation and no buyer has an incentive to deviate unilaterally.

There also exist 'exclusion equilibria' where the incumbent offers strictly positive compensations, at least to one buyer. These equilibria are such that $x_1 + x_2 \leq x_{\min}^b$, with $x_i > 0$ for at least one buyer. They are sustained by having both buyers rejecting the contract in the continuation equilibria following any offer where either $x'_1 < x_1$ or $x'_2 < x_2$. Hence, the incumbent would lose both buyers by decreasing the compensation and has no incentive to deviate.

Note that when offers can be discriminatory, the incumbent does not need to rely on coordination failures to exclude. Absent buyers' mis-coordination, the incumbent could sustain exclusion by offering x_{\min}^b to one buyer and zero to the other. However, the possibility to exploit coordination failures may make exclusion cheaper for the incumbent.

Finally, in a more general model with $N > 2$ buyers, one would obtain similar results if the following condition is satisfied: $N\pi^m(c_I) > N^*x_{\min}^b$, where N^* is the lowest number of buyers that the incumbent must sign in to make entry unprofitable. This condition says that fully compensating the minimum number of buyers such that entry is discouraged is profitable for the incumbent. If this condition is not satisfied, both 'entry equilibria' and 'exclusion equilibria' exist also when discriminatory offers are feasible, and the latter rely on coordination failures.

Sequential offers The neatest example of naked exclusion arises when the incumbent makes sequential offers. In our model, where $2\pi^m(c_I) > x_{\min}^b$, there exists a unique 'exclusion equilibrium' where the incumbent excludes offering zero compensations. To see why, let us analyse the second buyer's decision when the first buyer rejects exclusivity. If the second buyer also rejects, entry occurs and she ends up paying the competitive price c_I . If she accepts, she ends up paying the monopoly price. Hence, the second buyer requires at least x_{\min}^b to accept. For the incumbent it is profitable to have the second buyer sign behind the payment of (slightly more than) x_{\min}^b as entry is discouraged and

the incumbent earns $2\pi^m(c_I) > x_{\min}^b$. Hence, if the first buyer rejects, the second buyer signs. Let us consider now the case where the first buyer signed the contract. The second buyer anticipates that she will pay the monopoly price even though she rejects the contract, because her demand alone does not attract entry. Hence, the incumbent offers $x_2 = 0$ and the second buyer accepts. To sum up, the second buyer always accepts the contract, irrespective of the first buyer's decision. This implies that the first buyer will end up paying the monopoly price, both if she accepts and if she rejects the contract. Anticipating this, the first buyer is willing to accept even if $x_1 = 0$. Therefore, at this – *unique* – equilibrium, the incumbent offers $x_1 = 0$ to the first buyer, who accepts, and then $x_2 = 0$ to the second buyer, who also accepts. Pure exclusion occurs and it costs nothing to the incumbent.⁵⁰

3.4.3.2 Asymmetric buyers*

In this section we relax the assumption that buyers are identical. We assume, instead, that B_L , whose individual demand is given by $(1 + \alpha)D(p)$ is larger than B_S , whose individual demand is given by $(1 - \alpha)D(p)$. The parameter $\alpha \in [0, 1]$ measures the degree of asymmetry between the two buyers. Total market demand remains constant and is equal to $2D(p)$. The symmetric case reappears when $\alpha = 0$. The remaining assumptions are the same as in the game described in Section 3.4.3.1. In particular (as per assumption (3.10)):

$$c_I D(c_I) < f < 2c_I D(c_I).$$

In this setting, the scope for exclusion decreases as the asymmetry between buyers becomes more pronounced, as proven by the Proposition that follows. The analysis will focus on the case where the entry cost is sufficiently low, that is, $f < \frac{4}{3}c_I D(c_I)$. We will discuss at the end of the section the case where the entry cost is larger.

Proposition 3.1 *When buyers are asymmetric and the entry cost is sufficiently low, that is, $f < \frac{4}{3}c_I D(c_I)$, there exist two critical thresholds of the degree of asymmetry, α^* and α^{**} with $\alpha^* < \alpha^{**}$, such that:*

⁵⁰ If the condition $2\pi^m(c_I) > x_{\min}^b$ is not satisfied, there only exist equilibria where no buyers sign the exclusive contract.

- If the degree of asymmetry is limited ($\alpha \leq \alpha^*$), no buyer individually makes entry profitable. The results demonstrated in the symmetric case are still valid.
- If the degree of asymmetry is intermediate ($\alpha \in (\alpha^*, \alpha^{**}]$), then the large buyer's demand attracts entry and it is profitable for the incumbent to fully compensate the large buyer. It follows that:
 - when offers are simultaneous and homogeneous, there exist only 'entry equilibria';
 - when offers are either simultaneous and discriminatory or sequential, there exist a unique 'exclusionary equilibrium' where the incumbent fully compensates the large buyer and offers 0 to the small one.
- If the degree of asymmetry is pronounced ($\alpha > \alpha^{**}$), then the large buyer's demand attracts entry and it is not profitable for the incumbent to fully compensate the large buyer. Then, there exist only 'entry equilibria'.

Proof. The profits that the entrant makes by selling to the large buyer amount to $\pi_E(\alpha) = c_I(1 + \alpha)D(c_I)$. By assumption (3.10), when $\alpha = 0$ these profits are insufficient to attract entry, while when $\alpha = 1$ and the large buyer's demand represents total market demand, these profits suffice to. By continuity, $\pi_E(\alpha) \geq f$ iff $\alpha \geq \alpha^*$ where $\alpha^* = \frac{f}{c_I D(c_I)} - 1$.

Let us denote with $x_{min,L}^b(\alpha)$ and $x_{min,S}^b(\alpha)$ the compensation that makes a buyer – the large buyer and the small one, respectively – indifferent between paying the monopoly price and the competitive price. In this model with $D(p) = 1 - p$, $x_{min,L}^b(\alpha) = (1 + \alpha) \frac{3(1-c_I)^2}{8}$ and $x_{min,S}^b(\alpha) = (1 - \alpha) \frac{3(1-c_I)^2}{8}$. Let us also denote as $\pi_{TOT}^m = \frac{(1-c_I)^2}{2}$ the total monopoly profits the incumbent extracts from the two buyers. By (3.12) we know that when $\alpha = 0$ it is profitable for the incumbent to fully compensate one buyer: $\pi_{TOT}^m > x_{min}^b$; instead, by the monopoly deadweight loss, when $\alpha = 1$ it is not profitable to fully compensate the (large) buyer. By continuity, $\pi_{TOT}^m \geq x_{min,L}^b(\alpha)$, iff $\alpha \leq \alpha^{**} = \frac{1}{3}$. Note that the same threshold $\alpha \leq \alpha^{**} = \frac{1}{3}$ identifies also under which condition it is profitable for the incumbent to offer $x_{min,S}^b(\alpha)$ to both buyers: $\pi_{TOT}^m \geq 2x_{min,S}^b$, iff $\alpha \geq \alpha^{**}$. Finally, $\alpha^* \leq \alpha^{**}$ iff $f \leq \frac{4}{3}c_I D(c_I)$.

(i) If buyers' asymmetry is limited ($\alpha \leq \alpha^*$), no buyer is pivotal for entry. Following the same logic as in the symmetric case, one can easily demonstrate that similar results apply, with the following minor differences. When offers are simultaneous and homogeneous, 'entry equilibria' are such that $x \in [0, x_{min,S}^b(\alpha)]$: if the incumbent offered $x > x_{min,S}^b(\alpha)$, both buyers would accept (for the small buyer it is a dominant strategy to do so); however, from $\alpha^* < \alpha^{**}$ it follows that such an offer is not profitable for the incumbent. 'Exclusionary equilibria' are such that $x \in [0, \pi_{TOT}^m/2]$. When offers are simultaneous and discriminatory, there exist only 'exclusionary equilibria', with the incumbent offering $x_S + x_L \leq x_{min,S}^b(\alpha)$. Note that the highest compensation that the incumbent must offer to sustain exclusion is lower than in the symmetric case; also, the condition such that 'entry equilibria' do not exist (that is, $\pi_{TOT}^m > x_{min,S}^b(\alpha)$) is easier to be satisfied. In this respect asymmetry favours exclusion: since no buyer is pivotal, the incumbent can target the compensation to the small buyer, which makes exclusion cheaper. (ii) For intermediate degrees of asymmetry ($\alpha \in (\alpha^*, \alpha^{**}]$), the large buyer alone manages to attract entry. Then, coordination failures cannot be exploited to sustain exclusion because for the large buyer it is dominant strategy to reject exclusivity unless she is fully compensated. Hence, when offers are simultaneous and homogeneous, 'exclusionary equilibria' do not exist: to sustain exclusion the incumbent should offer $x_{min,L}^b(\alpha)$ to *both* buyers and this is not profitable. When offers are simultaneous and discriminatory, the incumbent can offer $x_{min,L}^b(\alpha)$ to the large buyer and 0 to the other. Since $\alpha \leq \alpha^{**}$ this offer is profitable for the incumbent. Note that, differently from the symmetric case, this equilibrium is unique. The reason is that the incumbent cannot take advantage of coordination failures to elicit both buyers' acceptance behind the payment of lower compensations. Also, the incumbent pays to the large buyer a compensation which is higher than the highest compensation paid in the symmetric case. The same equilibrium arises when offers are sequential. Note that when offers are sequential, under symmetry the incumbent can exclude at zero cost. (iii) When asymmetry is acute ($\alpha > \alpha^{**}$), the large buyer is pivotal for entry and it is not profitable for the incumbent to fully compensate her. Hence, 'exclusionary equilibria' cannot be sustained. ■

The new insight that arises when $f \geq \frac{4}{3}c_I D(c_I)$ is that, when the degree of asymmetry is intermediate, exclusion may be easier as compared to

the symmetric case. This situation arises when $\alpha^{**} < \alpha \leq \alpha^*$, with $\alpha^* > \alpha^{**}$ because the entry cost is large. When this is the case, no buyer is still pivotal for entry. However, the small buyer is sufficiently small to allow the incumbent to profitably offer $x_{min,S}^b(\alpha)$ to both buyers. As a consequence, ‘entry equilibria’ do not exist even in the case where offers are simultaneous and homogeneous. ‘Exclusionary equilibria’ arise and are characterised by different level of compensations depending on the behaviour of buyers in the continuation equilibria. Similarly, when offers are simultaneous and discriminatory, only ‘exclusionary equilibria’ exist. In the situation less favourable to the incumbent, the incumbent pays $x = x_{min,S}^b(\alpha)$ to the small buyer, which is smaller than the amount that the incumbent pays in the symmetric case. Finally, when offers are sequential, the incumbent manages to exclude at zero cost, but now the order followed to approach buyers matters for exclusion. Imagine the small buyer comes first. She anticipates that, if she rejects exclusivity, for the incumbent it is not profitable to fully compensate the large buyer (we are considering values of α such that $\alpha > \alpha^{**}$), so that the second (large) buyer will replicate her choice. Hence, the first buyer needs to be fully compensated to accept exclusivity. Since $\pi_{TOT}^m > x_{min,S}^b(\alpha)$, it is profitable for the incumbent to pay $x_{min,S}^b(\alpha)$ to the first buyer, and 0 to the second. Exclusion is achieved but it is costly for the incumbent. In contrast, imagine that the large buyer comes first. She anticipates that, if she rejects exclusivity, the incumbent will fully compensate the small buyer. Hence, the first (large) buyer is willing to accept exclusivity even for free. Now the incumbent excludes at zero cost. Hence, the incumbent will approach the large buyer first, if it can choose the order.

3.4.3.3 A large buyer with bargaining power*

In this section, we maintain the assumption that buyers are asymmetric: the individual demand of the large buyer B_L is given by $(1 + \alpha)D(p)$, while the demand of the small buyer B_S is given by $(1 - \alpha)D(p)$. We focus on the case of intermediate asymmetry, in which the demand of the large buyers is enough to make entry profitable, but the small buyer is not too small so that the monopoly profits extracted from her make it profitable for the incumbent to fully compensate the large buyer. In other words we assume that $f < 4c_I D(c_I)/3$ and we restrict the parameter α , that measures the degree of asymmetry between the buyers, to be included in $(\alpha^*, \alpha^{**}]$, with α^* and α^{**} being defined in the proof of Proposition 3.1.

In contrast to the previous section, we do not assume that the incumbent has all of the bargaining power. Rather, we assume that the incumbent makes a take-it-or-leave-it compensation offer to the buyer with probability β , and with probability $1 - \beta$ it is the large buyer that makes an offer regarding the level of the compensation, with $\beta \in [0, 1)$.

Proposition 3.2 *Assume that the entry cost is sufficiently low, that is, $f < \frac{4}{3}c_I D(c_I)$, and the degree of asymmetry between the two buyers is intermediate, that is, $\alpha \in (\alpha^*, \alpha^{**}]$. Then, when offers are either simultaneous and discriminatory or sequential, there exist a unique ‘exclusionary equilibrium’ where the incumbent and the large buyer agree on exclusivity and the large buyer receives a compensation $x_L(\beta, \alpha) = \beta \frac{3(1+\alpha)(1-c_I)^2}{8} + (1 - \beta) \frac{(1-c_I)^2}{2}$. The small buyer receives a compensation $x_S = 0$. The compensation received by the large buyer makes her strictly better off than in the case in which she rejects exclusivity and entry occurs: $x_L(\beta, \alpha) > x_{min,L}^b(\alpha)$.*

Proof. When the incumbent makes the offer, it offers the large buyer the compensation $x_{min,L}^b = \frac{3(1+\alpha)(1-c_I)^2}{8}$ that makes her indifferent between accepting and rejecting exclusivity. When, instead, it is the large buyer that makes the offer, it requires from the incumbent the compensation $x_{max,L}^b = \frac{(1-c_I)^2}{2}$: if exclusivity is not agreed upon, entry occurs and the incumbent makes 0 profits; under exclusivity, instead, entry is discouraged and the incumbent earns a total profits of $\frac{(1-c_I)^2}{2}$ from the two buyers. When she can make the offer, the large buyer appropriates those profits entirely. Hence, both in the case in which offers are simultaneous and discriminatory and in the case in which offers are sequential, the compensation that makes the large buyer agree on exclusivity is $x_L(\beta, \alpha) = \beta \frac{3(1+\alpha)(1-c_I)^2}{8} + (1 - \beta) \frac{(1-c_I)^2}{2} > x_{min,L}^b(\alpha)$ for $\beta < 1$. It is easy to see that the compensation $x_L(\beta, \alpha)$ is decreasing in β (that is, the stronger the large buyer’s bargaining power, the larger the compensation). ■

3.4.3.4 Downstream competition*

In this section, we assume that the good produced either by the incumbent or by the entrant is used by the two buyers as an input to produce a final good sold in a downstream market. For simplicity, we assume that there is a one-to-one relationship between the input bought by the buyer and the output sold in the final market, and that the cost of

transformation or resale is zero. The timing of the game is as follows. In the first stage the incumbent offers buyers a compensation x in order for her to accept exclusivity. Buyers decide. S denotes the number of buyers that accept exclusivity. In the second stage, after having observed S , the potential entrant decides on entry. In the third stage, firms active in the upstream market simultaneously set input prices. We restrict attention to linear price offers to buyers. The incumbent is able to set different wholesale prices to those buyers which have signed the exclusive contract and those that have not (free buyers). The potential entrant, if it has entered, can make offers to free buyers only. It offers a price w_E . In the last stage, buyers order the input and compete in the final market. Buyers do not have to pay any fixed cost to be active in the downstream market.

We will consider two extreme cases. First the case where downstream firms are independent monopolists, each facing demand: $D_i(p_i) = (1 - p_i)/2$ with $i = 1, 2$. Second, the case where downstream firms sell homogeneous products and compete à la Cournot, facing market demand $D(p) = 1 - p$.

In this setting, assumption (3.10) translates into:

$$\frac{c_I(1 - c_I)}{4} < f < \frac{c_I(1 - c_I)}{2}. \quad (3.13)$$

These restrictions on fixed costs ensure that, when downstream buyers are independent monopolists, entry sponsored by a single buyer is not profitable, whereas entry is profitable if E serves both buyers.

The remaining assumptions are the same as in Section 3.4.3.1.

Independent Monopolists

Proposition 3.3 *When downstream firms are independent monopolists:*

- (i) *there exist both ‘exclusion equilibria’ and ‘entry equilibria’, if the incumbent makes simultaneous and non-discriminatory offers;*
- (ii) *there exist only ‘exclusion equilibria’, if the incumbent makes simultaneous and discriminatory offers;*
- (iii) *there exists a unique ‘exclusion equilibrium’ where the incumbent excludes at no cost, if the incumbent makes sequential offers.*

Proof. Consider the three subgames that are relevant for the analysis. If both buyers accept the contract ($S = 2$), then the entrant does not enter ($\pi_{E|S=2} = -f$), and the incumbent sets the monopoly price $w_{I|S=2} =$

$(1 + c_I)/2$. Each buyer obtains a payoff $\pi_{B|S=2} = (1 - c_I)^2/32 + x_i$ and the incumbent earns $\pi_{I|S=2} = (1 - c_I)^2/8 - x_1 - x_2$.

The outcome is similar when a single buyer accepts exclusivity ($S = 1$). If it entered the market, firm E would set the price $w_{E|S=1} = c_I$ to the free buyer, that would sell the quantity $(1 - c_I)/4$ in the final market. The induced input demand is insufficient for the entrant to cover the fixed entry cost: $f > c_I(1 - c_I)/4$ by assumption (3.13). Then, following $S = 1$, entry does not take place and both buyers pay the monopoly price for the input, obtaining a payoff $\pi_{B|S=1} = \pi_{B|S=2} = (1 - c_I)^2/32$ (plus the compensation for the signer).

Finally, consider the case where both buyers reject the contract ($S = 0$). If the entrant enters the market, it will supply both buyers at the price $w_{E|S=0} = c_I$, obtaining a profit $c_I(1 - c_I)/2$. By assumption (3.13), this profit is sufficient to cover the fixed entry cost f . Hence, entry will follow $S = 0$ and the buyers' payoff will be $\pi_{B|S=0} = (1 - c_I)^2/8$.

Note that the compensation that makes a buyer indifferent between paying the monopoly price and the competitive price is given by:

$$x_{min} = \frac{(1 - c_I)^2}{8} - \frac{(1 - c_I)^2}{32} = \frac{3(1 - c_I)^2}{32}. \quad (3.14)$$

Also, paying the full compensation to both buyers is not profitable for the incumbent, as $\pi_{I|S=2} = \frac{(1 - c_I)^2}{8} < \frac{3(1 - c_I)^2}{16} = 2x_{min}$. However, it is profitable to pay the full compensation to a single buyer: $\pi_{I|S=2} = \frac{(1 - c_I)^2}{8} > \frac{3(1 - c_I)^2}{32} = x_{min}$.

Then, in this setting the same externality as in the basic model where buyers are final consumers arises and the same results are obtained. ■

Cournot Competition with Homogeneous Products We now study the case where downstream firms compete à la Cournot and sell homogeneous goods. We restrict the incumbent's marginal cost to be $c_I < 5/13$. The assumption that c_I is close enough to c_E implies that, in the sub-game following a single buyer accepting exclusivity, the entrant will serve the free buyer at the wholesale price c_I . One can check that for $c_I > 5/13$ the entrant will set $w_{E|S=1}^f < c_I$ at equilibrium and that our results are *a fortiori* valid. Further, this assumption is sufficient for both the signer and the free buyer to sell positive quantities at the equilibrium.

Proposition 3.4 *When downstream firms compete à la Cournot with homogeneous products, there exists a threshold level of the entry cost $f' = \frac{5c_I(1 - c_I)}{12}$ such that:*

- (i) if fixed costs are sufficiently low, that is, if $f < f'$, only 'entry equilibria' exist.
- (ii) otherwise, the standard result arises: 'exclusion equilibria' always exist. Also 'entry equilibria' exist when contract offers are simultaneous and non-discriminatory.

Proof. Let us analyse the three relevant subgames. If both buyers accept the contract ($S = 2$), then entry does not occur and the incumbent sets the monopoly price $w_{I|S=2} = (1 + c_I)/2$ to both buyers. The incumbent earns $\Pi_{I|S=2} = (1 - c_I)^2/6$ (gross of the compensations) and each buyer obtains a payoff $\pi_{B|S=2} = (1 - c_I)^2/36 + x_i$.

If both buyers reject the contract ($S = 0$) and entry takes place, the entrant's optimal price is $w_{E|S=0} = c_I$. The entrant's payoff amounts to $\pi_{E|S=0} = 2c_I(1 - c_I)/3 > f$ by assumption (3.13). Hence, entry is profitable when both buyers reject the exclusive contracts. In this case, buyers' payoff is $\pi_{B|S=0} = (1 - c_I)^2/9$.

Now, let us consider the case where a single buyer rejects the contract. This case is crucial to understand whether the externality central to the 'naked exclusion' framework arises also in this context. Imagine entry takes place. The entrant will supply the free buyer by charging the price c_I , while the incumbent will sell to the signer at the price $w_{I|S=1} = (1 + 3c_I)/4 > c_I$. Since the free buyer pays a lower price for the input than the signer, it will be a more efficient competitor downstream and will cover a larger share of the final market. The induced input demand for the entrant is thus larger than in the case of independent monopolists ($\frac{5(1-c_I)}{12} > \frac{1-c_I}{4}$) and the entrant's profits from the supply of a single buyer are also larger. As a consequence, when the fixed entry cost is not too large ($f < f'$), the demand of a single buyer is sufficient for the entrant to profitably enter the market. In such a case, the free buyer makes profits $\pi_{B|S=1}^f = 25(1 - c_I)^2/144$. The incumbent makes profits ($\Pi_{I|S=1} = (1 - c_I)^2/24 > 0$) because it sells a positive quantity to the signer (even though this is lower than the quantity sold by the entrant to the free buyer). The signer's profit amounts to $\Pi_{B|S=1}^s = (1 - c_I)^2/36$.

Then, the externality that is key for exclusion to take place does not arise in this case: the fact that a buyer accepts exclusivity does not force the other to accept as well. Indeed, since each buyer is pivotal for entry, in order to sustain an equilibrium where both buyers accept exclusivity, the incumbent should sufficiently compensate *both* buyers,

which makes exclusion unprofitable. For instance, when contract offers are simultaneous, the compensation that blocks unilateral deviations from $S = 2$ is $x_{min} = \pi_{B|S=1}^f - \Pi_{B|S=2} = 7(1 - c_I)^2/48$, and the incumbent's gain from exclusion is insufficient to profitably offer such compensation to both buyers: $\Pi_{I|S=2} = (1 - c_I)^2/6 < 2x_{min} = 7(1 - c_I)^2/24$.

The case where offers are sequential is slightly more involved but leads to the same conclusion. Let us start considering the case where *the first buyer signed the exclusive contract*. The second buyer anticipates that its input demand is sufficient to attract entry and requires $x_2 = \pi_{B|S=1}^f - \Pi_{B|S=2} = 7(1 - c_I)^2/48$ to sign. The incumbent cannot profitably induce this buyer to sign: the additional profits that it makes by excluding the rival is not large enough to pay such a compensation: $\Pi_{I|S=2} - \Pi_{I|S=1} = 3(1 - c_I)^2/24 < 7(1 - c_I)^2/48 = x_2$. Hence, if the first buyer signs, the second buyer rejects. (Note the difference: when the demand of a single buyer is insufficient to make entry profitable, the second buyer is willing to sign for free if the first buyer has already signed the exclusive contract.) Let us consider now the case where *the first buyer rejected the exclusive contract*. The second buyer anticipates that entry will occur irrespective of its decision to sign. Hence, it requires $x_2 = \pi_{B|S=0} - \Pi_{B|S=1}^s = (1 - c_I)^2/12$ to sign. Since entry occurs anyway, the incumbent's gain from having the second buyer sign the exclusive contract is $\Pi_{I|S=1} = (1 - c_I)^2/24 < x_2$. Therefore, it cannot profitably compensate the second buyer. Let us consider now the decision of the first buyer. Anticipating that the second buyer always rejects, the first one requires $x_1 = \pi_{B|S=0} - \Pi_{B|S=1}^s = (1 - c_I)^2/12$. But again, the incumbent's gain from having only one buyer sign the exclusive deal is insufficient to pay such a compensation. In equilibrium, both buyers reject the exclusive dealing offer and entry occurs.

When, instead, the fixed entry cost is large ($f \geq f'$), the demand of a single buyer is still insufficient to attract entry. Then we are back to the same situation as under independent markets where the incumbent can use exclusive contracts to exclude. ■

3.4.4 Penalty Fees for Exclusivity Breach*

In this section, we consider again the standard Chicago School setting with a single buyer, but we introduce the possibility for exclusive

contracts to include contractual damages, that is, penalty fees (agreed upon already at the contractual stage) payable whenever a buyer breaches the exclusivity clause and purchases the input from one of the incumbent's rivals (this analysis is originally due to Aghion and Bolton, 1987). We also allow the contract to include a contractual price, that is, an agreement on the price at which the good will be traded in later periods. Price commitment is not necessary for the results, but simplifies the exposition substantially. Other simplifying assumptions, yet immaterial for the results, are that the buyer wants to purchase at most one unit of the product, if the price is below her valuation $v = 1$, and that there are no fixed costs to enter the market.

We start the analysis from the case where the entrant's marginal cost is common knowledge at the contracting stage. We will then extend the analysis to the case where, at the contracting stage, the parties only know the distribution of the entrant's marginal cost.

3.4.4.1 Optimal contract absent uncertainty*

The timing of the game is the following. At time 1 the incumbent (whose marginal cost is $c_I < 1$) offers the buyer an exclusive contract (p, d) , where p is the price of the good that the incumbent commits to sell the good to the buyer at and d is the penalty fee that the buyer would need to pay to the incumbent if it dealt with E in spite of having signed up for exclusivity. At time 2 the buyer decides whether to accept such a contract. At time 3, after observing the buyer's decision, the entrant (whose marginal cost is $c_E < c_I$) decides whether to be active. At time 4, if an exclusive contract is in place, the entrant – if active – offers a price p_E to the buyer, who decides whether to breach. If there is no exclusive contract, active suppliers set the price. If firm E decided to enter, the suppliers compete à la Bertrand for the buyer.

Let us solve the game backwards. At time 4, absent exclusivity, the incumbent sets the monopoly price $p_I^m = 1$ if no entry occurred; if entry occurred, firm E supplies the buyer at a price $p = c_I$ and makes profits $\pi_E = c_I - c_E > 0$. If the exclusive contract is in place, and firm E decided to be active, the buyer breaches exclusivity if she is better off paying the price p_E and the penalty, rather than paying the contractual price p . Hence, the highest price that induces breaching is $p_E^b = p - d$.

At time 3 the entrant decides to enter the market both if the buyer rejected the exclusive contract and if the buyer agreed on a contract such that it is profitable to induce breaching (that is, such that $p - d \geq c_E$).

At time 2 the buyer decides on exclusivity. If she rejects the contract, entry will take place and she will enjoy surplus $1 - c_I$. If she accepts the contract, irrespective of whether breaching will occur, she will enjoy surplus $1 - p$. Then, she will accept the contract as long as the contractual price is (weakly) below c_I .

At time 1 the incumbent chooses the contract to offer. The optimal contractual price is the highest price among the ones that elicit the buyer's acceptance: $p^* = c_I$. If it offers a fully exclusive contract, that is, a contract with a high enough penalty fee to discourage breaching, the incumbent makes zero profits, as in the case where no contract is offered. Hence, a fully exclusionary contract is not profitable.⁵¹ A contract that allows for breaching must include a penalty fee $d \leq p^* - c_E = c_I - c_E$. The optimal contractual penalty thus is $d^* = c_I - c_E$. By offering this contract the incumbent accommodates entry and appropriates (through the penalty fee) the efficiency gains that the entrant brings into the market.

3.4.4.2 Optimal contract with uncertainty over the entrant's marginal cost*

In this section we consider the same setting as in the previous section, with the only difference that at the contracting stage parties do not know the exact value of the future entrant's marginal cost. They know its probability distribution though, and more specifically that the cost is uniformly distributed over the interval $[0, 1]$. The realisation of the entrant's marginal cost becomes common knowledge at time 3, when the entry decision is taken.

Let us start considering the case where no exclusive contract has been signed. At time 3 the entrant decides to enter the market if the realisation of its marginal cost c_E is below the incumbent's marginal cost c_I . In such a case, at time 4, it will win competition for the buyer setting the price c_I . If the realisation of the entrant's marginal cost is above c_I , entry does not take place and at time 4 the incumbent will set the monopoly price $p_I^m = v = 1$. Hence, absent the contract, the probability of entry amounts to $\phi = \text{prob}[c_E < c_I] = c_I$ and, at time 2, the buyer expects to enjoy surplus $(1 - c_I)c_I$ by rejecting exclusivity. The incumbent expects to make profits $\pi_I^{\text{noED}} = (1 - c_I)^2$.

⁵¹ This argument represents another way to illustrate the Chicago School critique.

Let us consider now the case where an exclusive contract (p, d) is in place. The entrant decides to enter the market if the realisation of its marginal cost is such that it could (profitably) offer a ‘competitive’ price to the buyer, namely that she would be happy to pay the penalty fee to I and buy from E at this ‘competitive’ price, which must be below $(p - d)$. The probability of entry is then $\phi' = \text{prob}[c_E < p - d] = p - d$. The buyer’s surplus, if she accepts the contract, is $1 - p$ irrespective of whether entry takes place.

At time 2 the buyer accepts an exclusive contract if the contractual price p is such that $1 - p \geq (1 - c_I)c_I$. Hence, at time 1 the contract that maximises the incumbent’s payoff will be identified by solving the following programme:

$$\max_{p,d} \left[\int_0^{p-d} (d) dc_E + \int_{p-d}^1 (p - c_I) dc_E \right] \text{ s.t. } p \leq 1 - (1 - c_I)c_I \quad (3.15)$$

It is easy to show that the solution is given by $(p^* = 1 - (1 - c_I)c_I, d^* = p^* - \frac{c_I}{2})$. Note that the optimal contract entails a probability of entry $\phi'^* = p^* - d^* = \frac{c_I}{2} < c_I$. Hence, under the exclusive contract, the probability of entry is below the efficient level $\phi = c_I$. Thus, compared to the earlier case of the entrant’s cost being deterministic, when this is uncertain, there is an excessive (that is, inefficient) level of exclusion.

Finally, note that the optimal contract allows the incumbent to make profits $\pi_I^{ED} = (1 - c_I)^2 + \frac{(c_I)^2}{4}$, which are larger than the profits obtained by the incumbent absent exclusivity.

3.4.5 Competing for Exclusivity*

We have so far considered the case where the incumbent has a first-mover advantage relative to the entrant. Let us next analyse a game where the incumbent and the entrant compete for exclusivity. The model is as close as possible to the other ones studied in Section 3.4.3 and in Chapters 1 and 2 (where exclusive dealing was not possible, though). As such, we are also back to a scenario that does not allow for the possibility for a contract to be breached or renegotiated.

As above, we assume that: demand for the homogeneous good is given by $p = 1 - Q$; the incumbent and the entrant have marginal costs characterised by $c_E = 0 < c_I < 1/2$; if it enters, the entrant has to pay a fixed cost f with $c_I(1 - c_I) < f < 2c_I(1 - c_I)$; there are two identical buyers B_1 and B_2 .

The game is as follows. In the first stage, the incumbent and the entrant simultaneously offer B_1 an exclusive contract including a compensation x_I^1 and x_E^1 , respectively. B_1 can either accept one of the exclusive dealing offers or reject both of them. In the second stage, firm E decides whether it wants to enter and if so it sinks its entry cost f . In the third stage, the two suppliers again compete in exclusivity for B_2 , who can either accept one of the two, or reject both. Finally, active firms set (linear) prices to buyers, consistently with the contracts that have been signed; for instance, a buyer who has signed exclusivity with firm i can buy only from that firm.

We prove the following:

Proposition 3.5 *There exist a threshold level of the incumbent's marginal cost, $\tilde{c}_I \equiv 1 - \sqrt{15}/5$, and a threshold level of the entry cost, $\tilde{f}(c_I) \equiv \frac{18c_I - 13c_I^2 - 2}{8}$, such that:*

- (i) *If either $c_I < \tilde{c}_I$ or $c_I \geq \tilde{c}_I$ and $f > \tilde{f}(c_I)$, then the first buyer accepts the exclusivity offer of the incumbent and the entrant does not enter the market. Both buyers pay the incumbent's monopoly price.*
- (ii) *If $c_I \geq \tilde{c}_I$ and $f \leq \tilde{f}(c_I)$, then the first buyer accepts the exclusivity offer of the entrant. The entrant enters the market. The first buyer pays the entrant's monopoly price, while the second buyer pays the competitive price c_I .*

Proof. We solve the model by backward induction. At the last stage, a buyer i who has accepted the exclusive dealing offer (henceforth, 'ED') of firm I pays $p_I^m = (1 + c_I)/2$ and obtains surplus $(1 - c_I)^2/8 + x_I^i$. Firm I has profits $(1 - c_I)^2/4 - x_I^i$. A buyer i who has accepted ED of firm E (and firm E has entered the market) pays $p_E^m = 1/2$ and obtains surplus $1/8 + x_E^i$. Firm E makes profits $1/4 - x_E^i$. If a buyer has rejected both ED and firm E did not enter the market, then at the last stage the buyer will pay $p_I = (1 + c_I)/2$ and obtain surplus $(1 - c_I)^2/8$. Firm I obtains profits $(1 - c_I)^2/4$ from this buyer. Finally, if a buyer has rejected both offers and the entrant has entered, then price competition results in the buyer buying from the entrant at price c_I and obtaining surplus $(1 - c_I)^2/2$. The entrant makes profits $c_I(1 - c_I)$ on this buyer.

Let us now move to stage 3. We have two cases to consider. (i) If firm E has entered, then whatever the outcome of the negotiations with B_1 , the second buyer rejects any ED offer at equilibrium. This is because, like in

the Chicago School setting, no firm is able to make a high enough offer to compensate B_2 of the lost surplus in case of acceptance. By signing with I , B_2 would lose $(1 - c_I)^2/2 - (1 - c_I)^2/8 = 3(1 - c_I)^2/8$, which is larger than the maximum amount that the incumbent is willing to offer, namely $(1 - c_I)^2/4$. By signing with E , B_2 would lose $(1 - c_I)^2/2 - 1/8$, while E would be willing to offer at most $1/4 - c_I(1 - c_I)$. It is easy to check that the latter is lower than the former under our assumption that $c_I < 1/2$. (ii) If firm E has not entered, then B_2 will have to pay monopoly price to I whatever she does. By indifference, she is willing to accept ED even at zero compensation.

At the entry stage, there are three possible cases. (i) If B_1 did not sign any ED, then firm E anticipates that upon entering it will serve both buyers at the duopoly prices, thereby making $2c_I(1 - c_I) - f > 0$: it enters. (ii) If B_1 signed with I , the entrant will not enter since it anticipates that it would make $c_I(1 - c_I) < f$. (iii) If B_1 had accepted E 's offer, then the entrant enters as it anticipates that the continuation profits $1/4 + c_I(1 - c_I) - f$ are positive. (Note that the entrant has already paid the compensation x_E^1 to the first buyer, hence the amount of the compensation does not affect the entry decision.)

Let us now turn to the first stage. If B_1 accepts ED from I , firm E will not enter. B_1 will pay p_I^m and its expected surplus will be $(1 - c_I)^2/8 + x_I^1$. If it accepts ED from E , then entry will occur. B_1 will pay the entrant's monopoly price and its surplus will be $1/8 + x_E^1$. By rejecting both ED offers, B_1 will pay the competitive price c_I and will have a surplus $(1 - c_I)^2/2$. Therefore, B_1 will prefer the incumbent's ED over the entrant's ED if (and only if) $x_I^1 \geq x_E^1 + 1/8 - (1 - c_I)^2/8$. Note that a compensation slightly higher than the entrant's is not enough for the incumbent to induce B_1 to prefer its own exclusivity offer over the rival's. The incumbent must also compensate the buyer for the loss it suffers paying its monopoly price rather than the entrant's lower monopoly price. Further, B_1 will prefer the incumbent's ED over rejection of both offers if (and only if) $x_I^1 \geq (1 - c_I)^2/2 - (1 - c_I)^2/8 = 3(1 - c_I)^2/8 \equiv x_{I,\min}^1$. Finally, B_1 will prefer the entrant's ED over rejection of both offers if (and only if) $x_E^1 \geq (1 - c_I)^2/2 - 1/8$.

What would be the maximum compensations that the sellers would be ready to offer to induce B_1 to accept the own exclusivity offer rather than the rival's? In the case of firm I , the incumbent knows that if B_1 signs an ED with it, then entry will not take place and it will make the monopoly profits in both periods, thereby obtaining $(1 - c_I)^2/2 - x_I^1$.

Instead, if B_1 accepts the entrant's exclusivity offer, entry will take place and the incumbent will make zero profits. Therefore, I 's maximum offer would be $x_{I,max}^1 = (1 - c_I)^2/2$.

Firm E knows that if B_1 accepts to sign an ED with it, then it will enter the market. It will make the monopoly profits on the first buyer and the duopoly profits on the second buyer. If instead the first buyer accepts the incumbent's exclusivity offer, then the entrant will decide not to enter and will make zero profits. Therefore, the highest offer firm E is willing to make will be $x_{E,max}^1 = 1/4 + c_I(1 - c_I) - f$.

Putting together the above considerations, the incumbent will be able to lure B_1 to exclusivity if and only if:

$$x_{I,max}^1 \geq \max \left\{ x_{E,max}^1 + \frac{1}{8} - \frac{(1 - c_I)^2}{8}, x_{I,min}^1 \right\}. \quad (3.16)$$

In other words, the incumbent must make an offer sufficiently large to outbid the entrant's offer and to induce the buyer to prefer exclusivity over rejection of both offers.

It is easy to check that $x_{I,max}^1 > x_{I,min}^1$. Also, $x_{I,max}^1 \geq x_{E,max}^1 + 1/8 - (1 - c_I)^2/8$ if (and only if) $f \geq \tilde{f}(c_I)$, with $\tilde{f}(c_I) < c_I(1 - c_I)$ for $c_I < \tilde{c}_I$. Hence when c_I is very low the threshold $\tilde{f}(c_I)$ is below the lowest admissible value of the entry costs, and the incumbent always wins competition for exclusivity. Note that, depending on the value of c_I and f , $x_{I,min}^1$ may be larger than the compensation $x_{E,max}^1 + 1/8 - (1 - c_I)^2/8$ that allows the incumbent to outbid the entrant. When this is the case, the compensation actually offered by the incumbent in equilibrium amounts to $x_{I,min}^1$.

When, instead, $x_{I,max}^1 \leq x_{E,max}^1 + 1/8 - (1 - c_I)^2/8$, it is firm E that will induce B_1 to accept the exclusivity by offering $x_E^{*1} = x_{I,max}^1 - 1/8 + (1 - c_I)^2/8$. (It is easy to check that $x_{I,max}^1 - 1/8 + (1 - c_I)^2/8 > x_{E,min}^1$, hence this offer allows firm E to induce the buyer to prefer its exclusivity offer over the rival's and over rejection of both offers.) ■

This section has proven that the incumbent can exclude a more efficient entrant by resorting to exclusive dealing also in a context in which the entrant can also bid for exclusivity. Similarly to the model of predation set out in Chapter 1, the reason why exclusion of a more efficient rival takes place is the following. Both firms anticipate that the seller that signs up the first buyer will eventually sell to the second. Therefore, each firm's bid for the first buyer depends on what is the

expected gain from serving the second buyer. While the incumbent extracts the monopoly profits if it serves the second buyer (because if it does not get the first buyer, the entrant will not enter), the entrant cannot get more than the duopoly profits (because the incumbent will compete in price for the second buyer). Only when the marginal cost of the incumbent is sufficiently large and the entry cost is not too significant, will the entrant be able to secure exclusivity with the first buyer.

Differently stated, exclusivity with the incumbent, by excluding the entrant and thus removing competition in the second period, allows larger rents to be extracted from the outsider of the first-period contracting. This increases the joint payoff of the incumbent, the entrant and the first buyer as compared to the situation where the first buyer accepts exclusivity with the entrant, and allows the incumbent to make an offer that the entrant is unable to replicate despite being more cost-efficient. Indeed, the condition that is established when the incumbent manages to outbid the entrant – $x_{I,max}^1 > x_{E,max}^1 + 1/8 - (1 - c_I)^2/8$ – can be given precisely this interpretation. If B_1 accepts exclusivity with the incumbent, the joint payoff of the incumbent, the entrant and the first buyer is given by:

$$\overbrace{\frac{(1 - c_I)^2}{4} + \frac{(1 - c_I)^2}{4}}^I + \overbrace{\frac{(1 - c_I)^2}{8}}^{B_1} + \overbrace{0}^E \quad (3.17)$$

If, instead, B_1 accepts exclusivity with the entrant, the joint payoff of the three agents is given by:

$$\overbrace{\frac{1}{4} + c_I(1 - c_I)}^E + \overbrace{\frac{1}{8}}^{B_1} + \overbrace{0}^I \quad (3.18)$$

The former is larger under the conditions indicated in the above Proposition.

Lemma 3.6 *At the exclusionary equilibrium, the incumbent is making a profit sacrifice on the first buyer.*

Proof. As we have seen above, when the exclusionary equilibrium exists, the equilibrium compensation paid by the incumbent to the first buyer

will be:

$$x_{I,max}^1 = \max \left\{ x_{E,max}^1 + \frac{1}{8} - \frac{(1-c_I)^2}{8}, x_{I,min}^1 \right\} \quad (3.19)$$

$$= \max \left\{ 1/4 + c_I(1-c_I) - f + \frac{1}{8} - \frac{(1-c_I)^2}{8}, 3(1-c_I)^2/8 \right\}. \quad (3.20)$$

Suppose $x_{I,max}^1 = 3(1-c_I)^2/8$ (which holds if $f > \widehat{f}(c_I) \equiv \frac{-12c_I^2+16c_I-1}{8}$). Since B_1 receives a compensation but pays the monopoly price, the profit made by the incumbent on B_1 will amount to: $\pi_I^1 = \frac{(1-c_I)^2}{4} - \frac{3(1-c_I)^2}{8} < 0$, that is, there would be a profit sacrifice.

Suppose instead that $x_{I,max}^1 = 1/4 + c_I(1-c_I) - f + \frac{1}{8} - \frac{(1-c_I)^2}{8}$, which holds when $f \leq \widehat{f}(c_I)$. In this case, the profit made by the incumbent on B_1 will be:

$$\pi_I^1 = \frac{(1-c_I)^2}{4} - \left[1/4 + c_I(1-c_I) - f + \frac{1}{8} - \frac{(1-c_I)^2}{8} \right] \quad (3.21)$$

We can check that $\pi_I^1 \leq 0$ iff $f \leq \frac{-11c_I^2+14c_I}{8} \equiv f_L(c_I)$. But recall that in the case we are studying here the condition $f \leq \widehat{f}(c_I)$ holds. It is easy to check that $f_L(c_I) \geq \widehat{f}(c_I)$, which implies that in the area we are considering it is always true that $f \leq f_L(c_I)$. Hence, the incumbent makes a loss on the first buyer. ■

3.4.6 ED and Investment Promotion*

In this section we study a model that illustrates the role of ED for investment promotion (or protection) and the link between investment promotion (or protection) and the risk of inefficient foreclosure.⁵²

The model The agents involved in the model are a buyer (B), an incumbent seller (I) and an alternative supplier (E). At date 0, B and I can sign an exclusive contract which prohibits B from trading with E . We assume that the initial contract is incomplete, in the sense that it cannot specify the terms of future trade because the nature of trade is hard to describe in advance. The only possible term in the initial contract, aside from a lump-side payment, is the exclusivity provision. In other

⁵² This model is largely based on Fumagalli et al. (2012).

words, the only fact that can be described *ex ante* and verified *ex post* is that *B* does not conduct any trade with another seller.⁵³ At date 1, if the contract is signed, but before trade, the incumbent may undertake non-contractible investment which affects the value of *ex post* trade.⁵⁴ Finally, at date 2 trade occurs.

Central to the analysis is the assumption that the exclusive contract cannot be renegotiated before trade takes place.⁵⁵ This is because we want to assess the welfare effects of ED in a setting where investment promotion and the risk of inefficient foreclosure can arise simultaneously.

For simplicity, we assume that the buyer demands at most one unit of a good. The two suppliers, *I* and *E*, are equally efficient (their marginal costs are normalised to zero: $c_E = c_I = 0$) but, absent investments, the buyer's valuation for *E*'s product is higher: $v_E > v_I$.

If *I* invests ψ into their relationship, which entails a cost $C(\psi) = (\eta\psi^2)/2$, with $\eta > 0$, *B*'s valuation for the incumbent's product becomes $v_I + \psi$.⁵⁶ However, the investment may also have a positive external effect, that is, it may increase the value of the relationship between the buyer and the alternative supplier which becomes $v_E + \lambda\psi$, with $\lambda \in [0, 1)$. Consider, for instance, the case of a manufacturer that invests in technical training of a retailer in order to improve the quality of the retailing service. Such an investment may also benefit the relationship between the retailer and rival manufacturers. When $\lambda = 0$, there is no external effect and investment increases only the internal value of the transaction.⁵⁷ Finally, we model pricing decisions at date 2 in the following way. With a probability 1/2 it is the buyer who makes the price offer, and with probability 1/2 it is the supplier(s) who makes (make) them.

⁵³ This form of contract incompleteness is typically assumed in the literature. This assumption, albeit extreme, captures the difficulty of contractually specifying all aspects of performance and allows us to study the effects of exclusivity and the interaction of foreclosure and investment promotion in the simplest possible setting.

⁵⁴ ED would not matter for investment promotion if investment could be specified in the initial contract.

⁵⁵ See Section 3.2.4.2 for a discussion of possible sources of renegotiation costs.

⁵⁶ We do not consider explicitly investment decisions by *E*. However, the valuation of *E*'s good can be thought of as determined by an entrant's exogenous investment.

⁵⁷ At the end of the section we will discuss the case $\lambda < 0$ as well as the case where the buyer is the investing agent.

3.4.6.1 When investment promotion facilitates foreclosure*

In this section we solve the game by backward induction and we start from the last stage where prices are decided for given contractual decision and investment decision.

Last stage payoff: the effect of ED for given investment

When the ED has been agreed upon, the buyer can trade only with the incumbent. When the incumbent makes the offer (which occurs with probability $1/2$), it charges the monopoly price $v_I + \psi$; when the buyer makes the offer, it requests the good for free. Hence, under exclusivity B and I share evenly the value of trade $v_I + \psi$. Firm E does not sell and its payoff is zero. The agents' payoffs (gross of investment costs) under exclusivity are indicated in Table 3.1 (top-left and bottom-left quadrants).

Table 3.1. Agents' payoffs with and without exclusivity

	ED	No ED
$v_I + \psi > v_E + \lambda\psi$	$\Pi_I = \frac{v_I + \psi}{2}$ $\Pi_B = \frac{v_I + \psi}{2}$ $\Pi_E = 0$	$\Pi_I = \frac{v_I + \psi - (v_E + \lambda\psi)}{2}$ $\Pi_B = \frac{v_I + \psi}{2} + \frac{v_E + \lambda\psi}{2}$ $\Pi_E = 0$
$v_I + \psi \leq v_E + \lambda\psi$	$\Pi_I = \frac{v_I + \psi}{2}$ $\Pi_B = \frac{v_I + \psi}{2}$ $\Pi_E = 0$	$\Pi_I = 0$ $\Pi_B = \frac{v_I + \psi}{2} + \frac{v_E + \lambda\psi}{2}$ $\Pi_E = \frac{v_E + \lambda\psi - (v_I + \psi)}{2}$

When no ED has been signed, the agents' payoffs depend on which good exhibits higher valuation after the investment. Let us start from the case where *ex post* E 's good is still valued more by the buyer (that is, $v_E + \lambda\psi \geq v_I + \psi$). With probability $1/2$ the buyer makes the offer, requires E 's good for free and appropriates the value of trade $v_E + \lambda\psi$. With probability $1/2$ suppliers compete. The pricing game is a standard asymmetric Bertrand game: the more efficient firm E wins the buyer offering a price $p_E = v_E + \lambda\psi - (v_I + \psi)$ and appropriates the additional value associated with its good. The argument is similar when it is I 's good that is valued more *ex post* (that is, when $v_I + \psi > v_E + \lambda\psi$), but it is the incumbent that appropriates the *ex post* efficiency gap. Table 3.1 also reports the agents' payoffs absent exclusivity (gross of investment costs).

Table 3.1 highlights that, for *given investment*, signing ED benefits the incumbent and harms both the buyer and the entrant. Note, however, an important distinction. When the incumbent is *ex post* more efficient

(that is, when $v_I + \psi > v_E + \lambda\psi$), ED redistributes welfare in favour of I but leaves total welfare unchanged. The buyer's loss is due to the fact that ED removes competition between the sellers, but trade nonetheless occurs with the more efficient supplier. Since firm E 's payoff is zero irrespective of exclusivity, the incumbent's gain coincides with the buyer's loss:

$$\Delta\Pi_I(\psi) = \frac{v_E + \lambda\psi}{2} = -\Delta\Pi_B(\psi) \quad (3.22)$$

Instead, when the entrant is more efficient *ex post*, by forcing trade with the incumbent ED forecloses the more efficient supplier and it not only redistributes total welfare but it *also reduces* it. In this case the loss caused to the buyer is larger than the incumbent's gain:

$$\Delta\Pi_I(\psi) = \frac{v_I + \psi}{2} < \frac{v_E + \lambda\psi}{2} = -\Delta\Pi_B(\psi) \quad (3.23)$$

The effect of ED on investment incentives

The above analysis shows that the exclusive contract will never be profitable for the buyer-incumbent coalition unless it affects the investment choice. We now investigate under which conditions this is the case.

Lemma 3.7 (ED and investment incentives)

When the ED is not renegotiable and the incumbent invests, there exists a threshold level of the investment cost, $\bar{\eta} = (1 - \lambda)^2 / 4(v_E - v_I)$, such that:

(i) if the investment cost is sufficiently large ($\eta \geq \bar{\eta}$), then exclusive dealing always promotes the investment: $\psi^{*ED} = 1/(2\eta) > 0 = \psi^{*NoED}$ for any level of the external effect $\lambda \in [0, 1)$;

(ii) if the investment cost is low enough ($\eta < \bar{\eta}$), then exclusive dealing promotes the investment if (and only if) the external effect is strictly positive ($\lambda > 0$): $\psi^{*ED} = 1/(2\eta) > (1 - \lambda)/(2\eta) = \psi^{*NoED}$; exclusive dealing is irrelevant for investment if there is no external effect ($\lambda = 0$)

Proof. Under exclusivity, the incumbent solves: $\max_{\psi} \left[\frac{v_I + \psi}{2} - \eta \frac{\psi^2}{2} \right]$.

The optimal level investment is identified by the FOC $\frac{1}{2} = \eta\psi^{*ED}$.

Absent exclusivity, the profit function of the incumbent is given by:

$$\pi_I(\psi) = \begin{cases} -\eta \frac{\psi^2}{2} & \text{if } v_I + \psi \leq v_E + \lambda\psi \\ \frac{v_I + \psi - v_E - \lambda\psi}{2} - \eta \frac{\psi^2}{2} & \text{otherwise} \end{cases}$$

Then, the level of investment which satisfies the FOC $\frac{1-\lambda}{2} = \eta\psi^*$ is globally optimal iff

$$\pi_I(\psi^*) = \frac{1}{2} \left(v_I + \frac{1-\lambda}{2\eta} - v_E - \lambda \frac{1-\lambda}{2\eta} \right) - \eta \frac{(1-\lambda)^2}{8\eta^2} > 0 \quad (3.24)$$

Condition (3.24) is satisfied iff

$$\eta < \frac{(1-\lambda)^2}{4(v_E - v_I)} \equiv \bar{\eta}.$$

Hence, absent ED, the optimal level of investment is $\psi^{*NoED} = (1 - \lambda)/2\eta$ when the investment cost is sufficiently low ($\eta < \bar{\eta}$). Notice that in this case the incumbent is made *ex post* more efficient: $v_I + \psi^{*NoED} > v_E + \lambda\psi^{*NoED}$. The optimal level of investment is $\psi^{*NoED} = 0$ otherwise. ■

To see the intuition, start from the situation where the investment cost is large enough ($\eta \geq \bar{\eta}$). Absent ED the incumbent does not invest. It anticipates that, in order to compete successfully with firm *E*, it should invest so much as to become *ex post* more efficient, but this is too costly. Investing less would not provide any benefit as post-investment competition results in the incumbent making no sales (see Table 3.1). Instead, under exclusivity, the incumbent anticipates that there will be no competition with the rival at date 2. Then, when it makes the offer, the incumbent will appropriate the value of trade with *B*, $v_I + \psi$, irrespective of the level of investment. Under exclusivity, the marginal benefit of the investment is thus larger and investment incentives are stronger. Note that, in this case, ED promotes investment even when there is no external effect ($\lambda = 0$), in contrast with Segal and Whinston's (2000b) 'Irrelevance Result'.⁵⁸

⁵⁸ The reason why the 'Irrelevance Result' does not hold in the model set out in this section is that ED is not renegotiable. This implies that the surplus shared *ex post*, under exclusivity, is $v_I + \psi$. Then, ED increases the incumbent's payoff by a term which responds to investment also when $\lambda = 0$. Instead, Segal and Whinston (2000b) assumes that (i) the initial contract is renegotiable so that when trade occurs, the parties are always able to appropriate the highest available surplus; (ii) the negotiation procedure is such that ED increases the incumbent's payoff by a term (the maximum surplus that the buyer and the entrant can generate) which does not depend on the investment when the external effect is absent. For this reason, ED has no impact on investment incentives when there is no external effect.

Consider now the case where the investment cost is sufficiently low ($\eta < \bar{\eta}$). Now, absent ED, investing so much as to become *ex post* more efficient is profitable for the incumbent. In such a case, post-investment competition results in the incumbent extracting its efficiency advantage ($v_I + \psi - v_E - \lambda\psi$), when it makes the offer. Under exclusivity, instead, post-investment competition is removed, and the incumbent will obtain the value of trade with the buyer, $v_I + \psi$, when it makes the offer. The comparison between the two payoffs highlights that, under exclusivity, investment is driven only by the internal effect: the incumbent takes into account only that additional investment increases the value of its trade with the buyer. Absent ED, instead, the incumbent takes into account that additional investment increases also the value of trade between the buyer and the rival supplier ($v_E + \lambda\psi$). Since $\lambda > 0$, the external effect is detrimental to the incumbent as it reduces the incumbent's *ex post* efficiency advantage. This suggests why, absent exclusivity, investment incentives are weaker as compared with the case where the ED is in place. When, instead, there is no external effect (that is, $\lambda = 0$), ED makes the incumbent earn a larger payoff but it does not affect the investment benefits at the margin, thereby leaving the equilibrium choice unchanged. In this case, the 'Irrelevance Result' does hold.

The contractual choice and welfare effects

At date 1, the incumbent and the buyer decide on exclusivity. Proposition 3.8 illustrates the equilibrium contractual choice and the welfare effects of ED.

Proposition 3.8 (Contractual choice and welfare effects of ED)

When the ED is not renegotiable and the incumbent invests, there exist two threshold levels of the investment cost, $\eta^s \equiv 3/4(v_E - v_I)$ and $\eta^w \equiv 3/8(v_E - v_I)$, with $\eta^s > \eta^w > \bar{\eta}$ for any $\lambda \in [0, 1)$, such that:

- (i) the ED is signed in equilibrium if (and only if) $\eta < \eta^s$;
- (ii) the ED is signed and is welfare-detrimental if (and only if) $\eta \in (\eta^w, \eta^s)$.

Proof. **Case (i)** Take the investment fixed at $\psi = \psi^{*NoED}$. Introducing exclusivity causes the incumbent a gain which is (weakly) lower than the

buyer's loss (see Table 3.1):⁵⁹

$$\Delta\pi_I(\psi^{*NoED}) = \pi_I^{ED}(\psi^{*NoED}) - \pi_I^{NoED}(\psi^{*NoED}) \quad (3.25)$$

$$\begin{aligned} &\leq -[\pi_B^{ED}(\psi^{*NoED}) - \pi_B^{NoED}(\psi^{*NoED})] \\ &= -\Delta\pi_B(\psi^{*NoED}) \end{aligned} \quad (3.26)$$

When $\eta \geq \bar{\eta}$, by Lemma 3.7, $\psi^{*NoED} = 0$. Since $v_I < v_E$, then $\Delta\pi_I(\psi^{*NoED}) < -\Delta\pi_B(\psi^{*NoED})$. However, by Lemma 3.7, ED also stimulates the investment ($\psi^{*ED} > \psi^{*NoED}$). By revealed preferences, this increases the incumbent's payoff under exclusivity as well as the buyer's payoff ($\pi_B^{ED}(\psi) = (v_I + \psi)/2$ is increasing in ψ), thereby expanding the incumbent's gain and mitigating the buyer's loss. The incumbent and the buyer jointly gain from the introduction of exclusivity when:

$$\frac{v_I + \psi^{*ED}}{2} - \eta \frac{(\psi^{*ED})^2}{2} > \frac{v_I + v_E}{2} - \frac{v_I + \psi^{*ED}}{2}$$

where $\psi^{*ED} = 1/(2\eta)$. This condition is satisfied iff the investment cost is low enough so that investment promotion is sufficiently strong: $\eta < \frac{3}{4(v_E - v_I)} \equiv \eta^s > \bar{\eta}$ for any $\lambda \in [0, 1]$.

When $\eta < \bar{\eta}$, by Lemma 3.7, ψ^{*NoED} is such that $v_I + \psi^{*NoED} > v_E + \lambda\psi^{*NoED}$. Hence, keeping the investment fixed at ψ^{*NoED} , $\Delta\pi_I(\psi^{*NoED}) = -\Delta\pi_B(\psi^{*NoED})$. When $\lambda = 0$ and ED leaves investment unchanged, the exclusive contract is signed by indifference. Instead, when $\lambda > 0$ ED promotes the investment, thereby making the incumbent's gain larger than the buyer's loss. Then the incumbent is always able to elicit acceptance in a profitable way:

$$\begin{aligned} \pi_I^{ED}(\psi^{*ED}) - \pi_I^{NoED}(\psi^{*ED}) &> \pi_I^{ED}(\psi^{*NoED}) - \pi_I^{NoED}(\psi^{*NoED}) \\ &= \Delta\pi_I(\psi^{*NoED}) \\ &= -\Delta\pi_B(\psi^{*NoED}) \\ &= \pi_B^{NoED}(\psi^{*NoED}) - \pi_B^{ED}(\psi^{*NoED}) \\ &> \pi_B^{NoED}(\psi^{*ED}) - \pi_B^{ED}(\psi^{*ED}). \end{aligned}$$

⁵⁹ We define as $\Delta\pi_i$ with $i = I, B$ the change of surplus of agent i due to the introduction of exclusivity.

Case (ii): When $\eta < \bar{\eta}$, firm E 's payoff is zero with and without the ED. Hence, total welfare coincides with B and I 's joint payoff and ED is welfare-beneficial whenever B and I have an incentive to sign it.

When $\eta \geq \bar{\eta}$, firm E is harmed by the introduction of exclusivity, because it supplies the buyer absent ED. Exclusive deal is welfare-detrimental iff investment promotion is not strong enough to make the increase in B and I 's joint payoff dominate the loss suffered by firm E :

$$\frac{v_I + \psi^{*ED}}{2} - \eta \frac{(\psi^{*ED})^2}{2} + \frac{v_I + \psi^{*ED}}{2} - \frac{v_I + v_E}{2} < \frac{v_E - v_I}{2} \quad (3.27)$$

Condition (3.27) is satisfied iff $\eta > \frac{3}{8(v_E - v_I)} \equiv \eta^w < \eta^s$ with $\eta^w > \bar{\eta}$ for any $\lambda > 0$. Hence, the ED is signed and is welfare-detrimental iff $\eta \in (\eta^w, \eta^s)$. ■

To see the intuition imagine there is no ED and consider the investment ψ^{*NoED} chosen by the incumbent. By Table 3.1 we know that introducing exclusivity – keeping the investment *fixed* at ψ^{*NoED} – benefits the incumbent but harms the buyer, and that the buyer's loss is (weakly) larger than the incumbent's gain: $\Delta\pi_I(\psi^{*NoED}) \leq -\Delta\pi_B(\psi^{*NoED})$.⁶⁰ Note that the previous inequality is strict when $\psi^{*NoED} = 0$ (that is, when $\eta \geq \bar{\eta}$) because firm E is more efficient and the introduction of ED forecloses its activity. It follows that, *absent any effect of ED on investment*, the Chicago School Critique applies: the lowest compensation that the buyer requires to sign the exclusive contracts is (weakly) larger than the incumbent's gain from having the contract signed; then the incumbent could never elicit the buyer's acceptance in a *profitable* way.

However, by Lemma 3.7 ED may stimulate the investment ($\psi^{*ED} > \psi^{*NoED}$). When this is the case, both the incumbent and the buyer's payoff under exclusivity increase (the former, by revealed preferences; the latter because higher investment increases the value of internal trade and the buyer appropriates part of this value), thereby *mitigating the buyer's loss* due to exclusivity and *expanding the incumbent's gain*. The lower the investment cost, the higher the investment increase spurred

⁶⁰ We denote with Π payoffs gross of investment costs and with π net payoffs. Since we keep investment fixed, the variation of the two due to the introduction of exclusivity is the same.

by ED, the more likely that the incumbent's gain becomes large enough to profitably compensate the buyer. Hence, if the investment cost is low enough, I and B jointly gain from the introduction of ED and have a private incentive to agree on it.

Let us now analyse the welfare effects of ED. When deciding on exclusivity, I and B do not internalise the effect of their decision on firm E and thus on total welfare. When the investment cost is low, that is, when $\eta < \bar{\eta}$, the incumbent invests even absent ED ($\psi^{*NoED} > 0$) and the investment makes I *ex post* more efficient than the rival supplier. In this case firm E will not supply the buyer even absent exclusivity, so that the introduction of ED does not affect its payoff. ED is signed in equilibrium and it is (weakly) welfare-beneficial. Instead, when the investment cost is large, that is, when $\eta \geq \bar{\eta}$, the incumbent does not invest absent exclusivity, and it is firm E that supplies the buyer ($v_E > v_I$). In this case, the introduction of exclusivity, by foreclosing its activity, harms firm E . Signing the contract will be welfare-detrimental when B and I 's joint gain from exclusivity is not large enough to compensate for the rival's loss. The higher the investment cost, the weaker investment promotion due to ED, the smaller the increase in B and I 's joint payoff, the less likely that this increase dominates firm E 's loss. It turns out that, when the investment cost is intermediate, investment promotion is strong enough to give B and I a private incentive to agree on exclusivity, but it is insufficient to make ED welfare-beneficial.

An alternative way to see why ED is welfare-detrimental is the following. Absent exclusivity, total welfare amounts to v_E because the incumbent does not invest and trade occurs with the more efficient firm E . ED, by forcing trade with the incumbent, promotes investment and increases the value of such trade – which becomes $v_I + \psi^{*ED}$. When the investment cost takes intermediate values this effect is sufficiently strong to make B and I willing to sign the contract; however, investment promotion is too weak to be beneficial for society because either (i) the incumbent remains less efficient than the rival ($v_I + \psi^{*ED} < v_E$); or (ii) the incumbent becomes more efficient than E , but achieving such an improvement is too costly for society. In both cases, ED is welfare-detrimental because B and I have a socially excessive incentive to use it.

The implication of the above Proposition is that an exclusive contract which forecloses a more efficient supplier may be signed *precisely* because it promotes investment. To see this, imagine that investment

is not possible – which in our setting corresponds to the case where the investment cost is prohibitively high. When $\eta \rightarrow \infty$ the ED is not signed in equilibrium: absent investment promotion there is no positive effect that can outweigh inefficient foreclosure so that the incumbent's gain from the introduction of ED is necessarily lower than the buyers' loss. Hence, foreclosure is not a concern in this case. Instead, when investment is possible and η takes intermediate values, the ED is signed (precisely because investment promotion induces *B* and *I* to agree on it), but is welfare-detrimental (because it leads to inefficient foreclosure or wasteful investment by *I*).

3.4.6.2 Discussion*

Negative external effect The above analysis assumes that the external effect of the investment is positive. However, an investment that increases the internal value may *deteriorate* the value of the external relationship. For instance, an investment that improves compatibility between a seller's input and the buyer's production process may make it more costly to use alternative inputs. When the incumbent invests and the external effect is negative, ED may *limit* rather than promote the investment. This is the case when the investment cost is sufficiently low ($\eta < \bar{\eta}$). As discussed above, absent ED, the incumbent decides investment taking into account also the external effect. When $\lambda < 0$, the latter is beneficial to the incumbent: by deteriorating the value of trade between the buyer and the rival supplier, additional investment increases the incumbent's *ex post* efficiency advantage. For this reason, absent exclusivity, investment incentives are stronger relative to the case where the ED is in place and investment is driven only by the internal effect. Lack of investment promotion implies that the buyer and the incumbent will not jointly benefit from ED and the contract will not be signed in equilibrium.

The buyer invests When the buyer invests in the relationship with the incumbent, the flavour of the results is similar to those illustrated above. ED is signed in equilibrium when the investment cost is low enough. For intermediate values of the investment cost, the signed contract is welfare-detrimental because it leads to inefficient foreclosure. There is a major difference though: when the buyer invests these effects arise only when the external effect is negative, because it is only in this case that ED promotes investment and will be signed in equilibrium.

The logic is again that, absent ED, the buyer takes into account both the internal and the external effect of the investment. When it is negative, the external effect is detrimental to the buyer. Hence, failing to internalise the external effect under ED makes investment incentives stronger.

The implication of the above analysis is that the identity of the investor and the sign of the external effect are crucial to determine when investment promotion may worsen rather than mitigate the concerns on the detrimental effect of ED. When the buyer invests, this is more likely to be the case if the external effect is negative. When the incumbent invests, this is more likely to be the case when the external effect is positive.

Single buyer *versus* multiple buyers The above analysis assumes that there exists a single buyer. Under multiple buyers, there are a number of situations where the incumbent is able to elicit acceptance on exclusivity even absent the investment. The key result from this section that investment promotion may facilitate inefficient foreclosure does not apply in this case. Instead, it holds when the incumbent is not able to. For instance: (i) when buyers are fierce downstream competitors; (ii) when buyer power is strong or there are weak economies of scale; (iii) when discrimination of contractual offers is not possible or not profitable.

3.5 From Theory to Practice

In this section, we discuss the policy implications that can be drawn from the literature reviewed above and we identify some criteria that can guide antitrust intervention concerning exclusive dealing contracts.

3.5.1 Differences between Exclusive Contracts and Rebates

We started this chapter giving emphasis to the fact that exclusive dealing contracts involve a *commitment by the buyer* not to purchase from alternative suppliers during a given reference period. This commitment component on the side of the buyer is not present in loyalty (or exclusivity) rebates, which are *unilateral* offers in which it is only the supplier that commits to offer different terms of trade depending on how much the buyer purchases. Hence, a buyer that enters an exclusive dealing contract with a supplier cannot purchase from another supplier. Instead, in the case of exclusivity discounts, a buyer can switch at any moment to an alternative supplier, even though it will obviously lose the discount. From the economic models analysed in this chapter, it emerged that the *ex ante*

commitment on the side of the buyer is a crucial factor for the incumbent's ability to use long-term contracts so as to deter efficient entry. Indeed, Section 3.2.3 has shown that, absent *ex ante* commitment on the side of the buyer, the incumbent would not manage to exploit its first-mover advantage so as to impede entry of the more efficient rival. Moreover, Section 3.2.5 has highlighted that lack of *ex ante* commitment prevents the incumbent from discouraging efficient entry also when there is no first-mover advantage and the rival can offer long-term exclusive contracts at the same time as the incumbent (the incumbency advantage still exists though, as in the model discussed in that section the rival is assumed not to have paid the entry cost yet when contract offers are made). Thus, even though it may appear that a discount conditional on buying 100 per cent (or most) of the buyer's requirement is equivalent to an exclusive dealing contract, the analysis developed in this chapter shows that the two differ in an important feature. The anti-competitive concern of exclusive dealing contracts is therefore more severe than the one of exclusivity discounts and of the other practices discussed in Chapters 1 and 2.

From the economic analysis of this chapter it also appears that exclusive dealing contracts not only make it more likely that the incumbent manages to exclude an as-efficient (or more efficient) rival, but also allow the incumbent to exclude at a lower 'cost'. In other words, a profit sacrifice is less likely to be a condition for exclusion to take place. Indeed, Section 3.2.3 has shown that when the incumbent enjoys a first-mover advantage, buyers are approached sequentially, and no buyer alone is large enough to attract entry, then the incumbent manages to secure buyers into exclusive dealing contracts without offering any compensation (or discount) to buyers and to extract full monopoly profits from them. Moreover, buyers' coordination failures may allow the incumbent to exploit its first-mover advantage and exclude without profit sacrifice also when exclusive dealing contracts are offered to all buyers simultaneously.⁶¹

These two considerations (*ex ante* commitment by a buyer and less likely need for a profit sacrifice by the incumbent) call for a *harsher treatment* of exclusive dealing contracts than both across-the-board price cuts which are suspected to be predatory pricing and rebate schemes (including loyalty rebates and individualised price discrimination). Furthermore, it does not

⁶¹ Note, however, that coordination failures may sustain exclusion without profit sacrifice also in settings in which there is no first-mover advantage and firms compete offering simple across-the-board prices (see Chapter 1, and in particular Sections 1.2.3.2 and 1.3.2.2) or rebate schemes (see Chapter 2, and in particular Sections 2.3.2.2 and 2.4.2).

seem reasonable to guarantee a ‘*price-above-cost*’ *safe harbour* to exclusive dealing contracts, differently from our discussion in Chapters 1 and 2 in relation to investigations involving predatory pricing or some type of discounting schemes.⁶²

The need to assess all relevant circumstances Nevertheless, an examination of the specific conditions offered to buyers in the exclusive dealing contracts can be highly informative. In particular, if it emerged that the profits extracted from buyers (the discounted sum of per-period profits, if the contract covers several periods) net of the compensation possibly offered upfront does not entail losses for the incumbent, it is important to ask why a rival, at least as efficient as the incumbent, is not able to secure orders from buyers. If there is scope for the as-efficient rival to make an offer which is more appealing without incurring losses, why is this not enough for the rival to be successful? Is the market share covered by the exclusive contracts sufficiently large to achieve exclusion? Is there scope for buyers’ coordination failures? Does the entrant suffer from the fact that buyers committed to (long-term) exclusive dealing contracts with the incumbent when the entrant was not in the position to make a counteroffer? Are the contracts already in place characterised by staggered expiry dates? Is the dominant firm in the position to induce buyers to accept the exclusive contract by imposing an ‘*all-or-nothing*’ *decision*? In other words, can the dominant firm threaten buyers not to supply any of their requirement in case they reject exclusivity? Is such a threat credible? (See below for a more extensive discussion of ‘*all-or-nothing*’ clauses.)

To sum up, as economists at least, we would find it difficult to find an infringement of competition law without a thorough analysis of the market and a careful understanding of why an as-efficient firm is not able to compete even if the incumbent does not make losses on *any subset* of buyers. Consistent with the messages from the models reviewed in this chapter, the analysis of whether the dominant firm suffers losses *should not be done on average*, across the contracts offered to all buyers, but at the level of a single buyer. From Section 3.2.3 we know that the dominant firm might suffer losses on specific buyers (large ones, buyers that are critical to facilitate or legitimise the entry of the rival, buyers who bear lower renegotiation costs), while recouping on others.

Similarly, we believe that any evidence that the incumbent is suffering losses on the contracts offered to some buyers should be complemented

⁶² See also the discussion in Fumagalli and Motta (2017a).

by a careful analysis of the case, which includes understanding whether there is a coherent story that rationalises an exclusionary objective served by exclusive dealing contracts (why are such buyers critical for entry or expansion? Why does securing such buyers into exclusive contracts discourage entry or impair the rival's ability to compete? Why such losses are likely to be recouped? Do the dominant firm and the rival compete for buyers' exclusivity?) and checking whether the facts of the case match the theory.

In both cases, *articulating a clear and consistent theory of harm* and checking that the facts of the case are consistent with the proposed theory should be of primary importance. Moreover, the process of assessing whether the facts of the case are consistent with the proposed theory of harm may also involve the analysis of the *effects* of the alleged exclusionary conduct on the competitors of the dominant firm.

This chapter has offered a number of hints as to possible theories of exclusion through exclusive dealing contracts, and to possible factors that should be looked at during an investigation. We summarise them in what follows.

3.5.2 Theory of Harm and Factors which Facilitate Exclusion

There are a number of conditions that are required for the theories of harm from exclusive contracts reviewed in this chapter to apply, as well as factors which may facilitate exclusion based on such theories of harm.

To start with, successful entry (or expansion) needs to be patronised by a sufficient number of buyers. This may be the case when entry involves important fixed costs of operation (relative to the total size of the market) or more generally when there are *significant scale economies* from the supply side, so that the entrant needs to sell to a sufficient number of buyers in order to achieve the efficient scale. This may also be the case when scale economies arise from the demand side, for example due to the presence of network externalities, and the entrant needs to reach a critical mass of users for its product to gain appeal in the eyes of consumers. Recall also that merely observing that the rival is already active in the marketplace is not sufficient to conclude that scale economies do not play a role any longer and then that exclusive contracts cannot exert adverse effects. Indeed, scale may be crucial for the rival to recover investment in additional capacity. Then, exclusive contracts might be used to prevent the expansion of smaller rivals already in the market. Alternatively, exclusivity involving buyers/retailers that are able to legitimise the quality of the entrant rather

than being aimed at deterring entry may have the purpose of delaying the resolution of consumers' uncertainty about the entrant's product, thereby limiting the expansion of the rival. This implies that one should check whether the exclusive contract is actually aimed at preventing the rival from achieving efficient scale. This is less likely to be the case if the *coverage* of exclusive contracts is small, that is, the share of total buyers in the market under exclusivity with the dominant firm is limited and if the *duration* of the exclusivity obligation is short. By contrast if the proportion of buyers involved in the exclusive contracts is large and the duration of exclusivity is long it is more likely that they can be used to exclude the rival. However, contracts of longer duration might be more effective in promoting relation-specific investments, if there exist the conditions for investment promotion to arise (see below). Also this element should be taken into account before concluding that longer exclusive contracts raise more severe anti-competitive concerns.

Moreover, there must be deep asymmetries between the incumbent and the entrant (or smaller rivals), which may be proxied by asymmetries in the investment in a crucial infrastructure and/or by a high and persistent market share of the incumbent. In other words, the existence of a *dominant position* is a necessary requirement for the finding of abusive exclusive contracts also from an economic perspective, and the stronger the dominance the more severe the anti-competitive concern should be, all else equal. The incumbent's dominant position may be reinforced by a *first-mover advantage* that arises when the incumbent is able to offer long-term contracts to buyers before the rival can react and make a counter-offer. This may occur because rival firms are not yet credible suppliers and then are not in a position to act as an effective constraint when buyers decide whether to accept an exclusive contract with the incumbent. The theory has shown that inefficient foreclosure is more likely when the incumbent can exploit such a first-mover advantage. However, we showed in Section 3.2.5 that exclusive dealing may lead to inefficient exclusion also when the first-mover advantage hypothesis is removed. In other words, the existence of a superior market position may be sufficient for exclusive dealing to have anti-competitive effects even when the dominant firm and its rival can compete for buyers' exclusivity, in parallel to what we show in the other chapters of the book concerning predatory pricing, rebates and margin squeeze.

Exclusive contracts are more likely to raise anti-competitive concerns when there are many buyers *whose individual demand is insufficient to sponsor entry* (or, more generally, to allow the rival to achieve

efficient scale) and who are unable to coordinate their purchasing decisions. In contrast, buyer concentration (that is, the existence of few large buyers) may significantly alleviate (and potentially eliminate) the anti-competitive concerns of exclusive contracts. In this respect, central purchasing agencies, by grouping orders of affiliated buyers, may create sufficient buyer concentration as to avoid inefficient exclusion and may also mitigate the risk of coordination failures. Moreover, the presence of few large buyers facilitates renegotiation of the initial contract with the incumbent, thereby hindering inefficient exclusion through this additional channel.

The economic literature discussed in this chapter has also shown that the risk of inefficient foreclosure is particularly high when the dominant firm can target specific buyers, that is, it can *discriminate contractual conditions* across different buyers and implement a divide-and-conquer offer, whereby it fully compensates in return for exclusivity a sufficiently high number of buyers (such that the remaining ones are insufficient to make entry profitable) and suffers losses on them, while extracting the monopoly rents from the others.⁶³

So far we have referred to a setting in which asymmetries across buyers are limited: we have contrasted the case in which there are many small buyers, whose individual demand is insufficient to sponsor entry, with the case in which there are few large buyers, whose individual demand is instead large enough to trigger entry. Let us consider now the case in which asymmetries are more pronounced and few large buyers, that are big enough to promote entry, coexist with small buyers. In this case large buyers are *critical* for entry. More generally, critical buyers may be buyers that have the ability to legitimise the entrant in the eyes of consumers or buyers whose contribution for the success of entry is particularly important for reasons other than size. Sections 3.2.3.1 and 3.4.3.2 show that *the presence of critical buyers makes exclusion less likely and more costly*. When contract offers are simultaneous, the dominant firm cannot take advantage of buyers' coordination failures so as to exclude with zero or limited compensations. Then, when contract offers need to be homogeneous, exclusion is no longer possible. When contract offers can be discriminated, exclusion is possible but the incumbent must compensate the critical buyers fully. Similarly, when offers are sequential, the incumbent no

⁶³ Of course exclusion is a profitable strategy as long as the monopoly rents extracted from the other buyers are large enough to compensate for the losses suffered on buyers that are fully compensated.

longer manages to induce all the buyers to accept exclusivity with zero compensations, but it must fully compensate the critical buyers.⁶⁴

We can draw two implications from this analysis. First, when there are critical buyers the ability to discriminate contractual conditions does not simply facilitate exclusion, but it is necessary for exclusion to take place. Second, when there are critical buyers, exclusive dealing contracts cannot lead to exclusion *without a profit sacrifice (on some buyers)*. In turn this implies that for exclusion to be a profitable strategy, the losses suffered on these critical buyers need to be compensated by the rents extracted on the others. This is likely to be impossible when asymmetries across buyers are *very pronounced*: in order to sustain exclusion the incumbent should fully compensate any large (or critical) buyer, but this is likely to be unprofitable because the rents extracted from small buyers are likely to be insufficient to cover the losses suffered on large ones. For instance, in the *Ticketmaster* case in Ireland (see Section 3.7.8), one might argue that Ticketmaster's exclusive contracts with event promoters were not anti-competitive because the two largest customers accounted for the vast majority of the tickets sold by Ticketmaster between 1998 and 2004. Given these features, it is unlikely for the dominant firm to be able to compensate the large customers (or buyers) at the same time as extracting sufficient monopoly rents from the rest of the market.

Note that large buyers (or more generally critical buyers) may strictly benefit from agreeing on exclusivity. Indeed exclusivity, by limiting upstream competition, allows the 'coalition' formed by the incumbent and the large buyers to extract monopoly rents from the small buyers; when the large buyers have some bargaining power, they appropriate a sufficiently large part of those rents to find exclusivity strictly profitable. Since the benefits of exclusion accrue to both the incumbent and the

⁶⁴ Note that, when there are large (or critical) buyers, exclusive dealing contracts can lead to exclusion as long as the incumbent enjoys a first-mover advantage. When, instead, the dominant firm and the rival compete for exclusivity the presence of large buyers essentially eliminates the exclusionary role of exclusive dealing contracts. In particular, if the competing firms make contractual offers simultaneously to all buyers, exclusion relies on buyers' coordination failures which cannot be exploited by the dominant firm when there are large buyers. When buyers are approached sequentially, and the large buyer comes second, exclusion cannot arise because rents extraction from the second buyer favourable to the incumbent does not take place. It is only when the large buyer comes first that exclusion is conceivable but unlikely. Consistently, in the models developed in Chapters 1 and 2 to study predation and rebates, the dominant firm and the rival make price offers to buyers at the same time and the existence of large buyers essentially eliminates the anti-competitive concern.

large buyers, it appears that exclusive dealing should be considered not only as a potential abuse of dominant position, but also as a potential *anti-competitive agreement*. Whereas in general small fragmented buyers could not be held responsible for accepting exclusive dealing contracts which have the effect of excluding some suppliers, the same cannot be argued for large buyers that alone can command a very large share of the demand, and that have received significant compensations for signing the agreement.

The economic literature has shown that the fact that *buyers are approached sequentially* makes exclusion more likely. In addition, the *staggering of contract expiry dates* that occurs if contracting is sequential can ensure that the number of buyers that are up for renewal at any point in time is low (assuming contracts of limited duration), thereby allowing the dominant firm to preserve exclusion over time. In Section 3.7.6, we discuss the *Nielsen* case, where a Canadian Tribunal considered Nielsen's practice of strategically staggering contract renewals as an element reinforcing the exclusionary purpose of exclusive contracts.

Sections 3.2.3.2 and 3.2.2 have also highlighted that the intensity of both *downstream and upstream competition* matters for the anti-competitive effect of exclusive dealing. Exclusive contracts are more likely to be exclusionary when downstream competition is particularly weak. Whether, instead, fierce downstream competition facilitates a dominant firm in excluding a more efficient rival depends on specific features of the market that need to be evaluated case-by-case. It is hard to make a general statement in this respect. Further, the risk of foreclosure is higher when post-entry competition between upstream firms is expected to be weak. We note, however, that (as set out in previous chapters as well) the degree of competition is not a variable which can be easily measured.

Finally, it is more likely that exclusive contracts lead to inefficient exclusion if the dominant firm is in a position to impose an '*all-or-nothing*' clause, that is, if the dominant firm can threaten buyers not to supply any of their requirement in case they reject exclusivity. This threat may be particularly effective when smaller rivals are unable to credibly contest a significant proportion of the demand of the dominant firm, due for example to capacity constraints, or reputational concerns. For instance, in the *Dentsply* case (reviewed in Section 3.7.5), the dominant supplier (with a market share of about 75–80 per cent) offered a full line of varieties of artificial teeth, including well-known 'must-have' brands, while its smaller rivals were offering a much restricted set of varieties. In similar situations, forcing an 'all-or-nothing' decision limits considerably the buyers' benefit

from rejecting exclusivity and makes it cheaper for the incumbent to elicit acceptance. Under those circumstances, also exclusive contracts of short duration (indeed terminable at will as in the *Dentsply* case) may lead to anti-competitive effects. Note that an ‘all-or-nothing’ clause suffers from a severe *credibility* problem: if a buyer rejects the exclusive contract, the dominant firm has no interest in refraining from ultimately supplying the buyer for the requirements that smaller rivals leave unsatisfied. Hence, the dominant firm needs to have a reputation for enforcing an ‘all-or-nothing’ clause in order to make it an effective instrument to foreclose rivals, a reputation possibly built during past interactions with buyers. Then, in order to consider such clauses as an element which makes the risk of foreclosure more likely, their credibility must be carefully assessed. For instance, as we will discuss in Section 3.7.5, the records in the *Dentsply* case contain examples of such threat being carried out in previous periods.⁶⁵

3.5.3 Pro-competitive Effects of Exclusive Dealing

So far we have discussed the anti-competitive role of exclusive contracts. However, one should try to understand whether behind exclusive contracts there are objective justifications or efficiency rationales. The economic literature has emphasised that exclusive contracts can also generate important efficiency gains. As discussed in Section 3.3.1, exclusive contracts may promote efficient investment if the investment is *relation-specific* and *non-contractible*. Under these requirements, there exists scope for opportunistic behaviour going against the investing agent once the investment has been undertaken, which *ex ante* may lead to underinvestment (a ‘hold-up’ problem). Exclusive contracts may then be necessary to restore efficiency.⁶⁶

Relation-specific investment means that once it has been incurred, the investment cannot be used to improve the value of trade with another buyer or another seller. *Non-contractible investment* means that the investment is either not observable or observable but not verifiable. As a consequence of these features, the investment cannot be measured by third parties like a court. Then, it is not possible to enforce a contract which establishes payments contingent on the level of investment which actually takes place after the contract is signed, ensuring that the investing

⁶⁵ The credibility of the threat made by the dominant firm is also central to the discussion of the *Meritor* and *Eisai* rebates cases analysed in Chapter 2.

⁶⁶ Section 3.3.1 provides several examples of investment where both characteristics are present.

party is compensated appropriately for the investment undertaken. If investment is contractible, and such contracts can be written, the efficient level of investment will be achieved with no need to rely on exclusive contracts.

Competition authorities should carefully verify the existence of both characteristics before accepting investment promotion arguments as a justification for exclusive dealing. Another reason for a careful approach towards investment promotion justifications is that the effectiveness of exclusive contracts in this respect depends crucially on the *sign of the external effect* – that is, on the way the investment affects the value of trade with other buyers or sellers, which might be detrimental or beneficial to the investing agent – and on the *identity of the investor* – whether it is the buyer or the seller. As discussed in Section 3.3.1, when the seller is the investing agent exclusive contracts are more likely to stimulate investment if the external effect increases the value of trade between the buyer and other sellers. *Vice versa*, when the buyer is the investing agent, exclusive contracts are more likely to promote investment if the external effect deteriorates the value of trade between the buyer and other sellers. These two features should be assessed with care before concluding that exclusive contracts are welfare-beneficial because of investment promotion.

Finally, the economic analysis suggests that the balancing exercise of any anti- and pro-competitive effects of exclusive dealing might be even more complex than already acknowledged, since the very existence of investment-promotion effects – rather than being only efficiency-enhancing – might actually make it easier for exclusive dealing to foreclose efficient entry.

All of the above considerations point to the conclusion that the burden of proving the pro-competitive effects should be on the defendant and that the stronger the defendant's dominant position, the stronger the proven efficiency gains should be.

3.6 Case-law

Exclusive dealing is another area where we find a major divide in the case-law across the Atlantic. In Europe, a formalistic approach has prevailed (with some minor exceptions we discuss below). In the US, by contrast, the focus has moved from a formalistic assessment of the behaviour in question towards the need to show likely anti-competitive effects arising from upstream foreclosure (for a finding of liability).

3.6.1 United States

The two early landmark judgments in the US Supreme Court jurisprudence on exclusive dealing came both in 1922.⁶⁷ In *Standard Fashion*, a producer of dress patterns (Standard Fashion) illegally engaged into an exclusivity agreement with Magrane-Houston, a Boston-based dry-goods store.⁶⁸ At the time, Standard Fashion was found to have a national market share of 40 per cent, with exclusive dealing arrangements with all the dress pattern outlets it served across the US. The Court was particularly concerned about small communities being served by a monopoly manufacturer. Moreover, it was worried about the exclusionary trend that these agreements could bring about: the 40 per cent exclusivity coverage may have quickly grown to 100 per cent (the Court did not substantiate its fear). Posner (2001: 251ff) commented that Standard Fashion's behaviour was likely to be exclusionary in light of the economies of scale in distribution: department stores were not common in the early 1920s, especially in smaller towns, so entry by firms offering single-product lines (and distributing via such malls) was unlikely; alternatively, entering with a full-product line and own retail stores would have been both risky and very expensive. Thus, Posner continued, Standard Fashion's strategy may have been costly in the short term, yet worthwhile in deterring entry and retaining supra-competitive profits for some time.

In *United Shoe Machinery*, the Supreme Court held that a series of practices by a near-monopolist amounted *de facto* to anti-competitive exclusive dealing. United Shoe Machinery was found to have over 95 per cent of US market for the supply of machinery used in shoe manufacturing. It sold such machinery to manufacturers under a lease system with various restrictive features, including the tying of input supplies and machinery and exclusivity clauses. The Court argued that, considering United Shoe Machinery's dominance, such restrictive practices

effectually prevent[ed] [the lessee] from acquiring the machinery of a competitor of the lessor except at the risk of forfeiting the right to use the machines furnished by [United Shoe Machinery], which may be absolutely essential to the prosecution and success of his business. (Para. 458 of the judgment)

The Court, however, did not provide the reason behind such essentiality.

⁶⁷ In this section we draw from the works of legal scholars such as Posner (2001), Bhatia et al. (2006), Gavil et al. (2008). The reader is also directed to Jacobson (2002) for an excellent review of exclusive dealing cases in the US, including several from the beginning of the twentieth century.

⁶⁸ Here, we also draw from Marvel (1982).

Over two decades later, in the late 1940s, the Supreme Court, in *Standard Stations*, set a rule of reason based on a test assessing whether an exclusive dealing agreement forecloses a 'substantial share of the line of commerce affected' (para. 314). In that case the Supreme Court found that Standard Oil – holding a market share of only 23 per cent – had illegally engaged in exclusive dealing with independent gas outlets that accounted for only 6.7 per cent of total gas sales in the area. Exclusive arrangements were also found to be widely used by Standard Oil's top six competitors. The Court recognised the potential for pro-competitive effects arising from such agreements but ruled that there would be liability as soon as competition was foreclosed in a *substantial* share of the market.

The next phase in the US antitrust treatment of exclusive dealing, several commentators noted (for example, Jacobson, 2002, and Gavil et al., 2008), began in 1961 with the Supreme Court ruling in *Tampa*. There, the Court found no antitrust infringement. Tampa Electric, a public utility, had signed a 20-year exclusive agreement with a coal mining company. After a disputed market definition, the Court found that the contract foreclosed less than 1 per cent of the relevant coal market. With *Tampa*, arguably, the Court adopted a less mechanical approach to exclusive dealing arrangements, or what we could call a more 'refined' rule of reason. The defendant's actual market power would be taken into account when assessing the potential anti-competitive effects, just as the contract duration, the contract notice period, the proportion of the market foreclosed by such agreements, whether competitors also used such arrangements, whether competitors could reach final consumers via alternative distribution channels and whether there were substantial barriers to entry.

This more tolerant approach towards exclusive dealing led to firms with more market power being cleared under US antitrust law, even when exclusive agreements covered larger proportions of the market. Two oft-cited cases from the late 1990s are *Omega Environmental* and *CDC*. In *Omega Environmental*, an Appeals Court held that exclusive dealing agreements with 38 per cent of the distributors in a market did not cause anti-competitive effects. The product in question was petrol (gas) dispensing equipment, to be used by retail petrol stations. The defendant (Gilbarco) had a 55 per cent share of the US market; it sold about 70 per cent of its dispensing equipment via distributors and the remaining part directly. The Court's reasoning was based on the short contract duration (with a typical notice period of 60 days), on the availability of alternative distribution channels (direct sales and other distributors) and on low

barriers to entry and expansion, even in the presence of exclusive dealing contracts.

In *CDC*, the ‘substantiality’ test was further stretched: an Appeals Court affirmed a lower court’s grant of summary judgment for the defendant (IDEXX), concluding that, in spite of its 80 per cent market share, its exclusive agreements with 65 per cent of the distributors did not foreclose rivals nor increased prices. The market concerned was blood analysis machines to veterinarians. IDEXX was actually a later entrant into the market and managed to engage into exclusive agreements with CDC’s legacy distributors. The Appeals Court agreed with the arguments raised by the lower court that direct dealing was feasible and that distributors provided limited additional services. Further, it pointed out that the exclusive agreements could be terminated on short notice (60 days) and that barriers to entry into distribution were low.

The tide turned in *Microsoft*, which involved lengthy and complex proceedings. Here, we focus on the part of the Appeals Court’s judgment relevant for exclusive dealing.⁶⁹ The DOJ had challenged Microsoft’s agreements with Internet access providers (IAPs). In particular, according to the DOJ:

Microsoft agreed to provide easy access to IAPs’ services from the Windows desktop in return for the IAPs’ agreement to promote [Microsoft Internet Explorer] exclusively and to keep shipments of internet access software using Navigator [a competing browser] under a specific percentage, typically 25%. (Para. 108 of the judgment⁷⁰)

The Appeals Court held that Microsoft’s exclusive contracts with 14 of the top 15 IAPs in North America foreclosed a substantial percentage of a key distribution channel for browsers (internet downloads or retail computer outlets were deemed less effective alternatives). In turn, this kept the adoption of Netscape Navigator (and other rival browsers) below a critical level, thwarting Netscape’s ability to challenge Microsoft in the operating systems’ market. Moreover, the Court could see no pro-competitive justification for this behaviour (Microsoft argued *inter alia* that it wanted to keep IAPs focused on Microsoft’s own products).

⁶⁹ We also draw from Motta (2004: section 7.5.2.2). We discuss other cases involving Microsoft later in this book: see Chapter 4 (in particular, Sections 4.6 and 4.7) for tying aspects of such cases; and Chapter 5 (in particular, Sections 5.5 and 5.6) for interoperability aspects.

⁷⁰ At para. 111 of the judgment, the Court also mentioned that this proportion amounted to 15 per cent in the case of AOL, the leading IAP at the time.

Another notable case was *Visa*, where an Appeals Court affirmed a lower court's judgment condemning rules set by Visa and MasterCard which prohibited their member banks from issuing payment cards through competing networks (that is, Discover Card and American Express). Visa and MasterCard were found to have, at the time of the case, market shares of 47 per cent and 26 per cent, respectively. The Court argued that, as a result of the exclusivity arrangement, Visa and MasterCard could raise their interchange service fees, as well as limit innovation. The defendants responded that their competitors (American Express and Discover) were successful in distributing their cards directly (that is, not via banks), so that final consumers were not affected by the (upstream) exclusive agreements, but the Court rejected this argument.

Another important case on exclusive dealing was *Dentsply*, decided in 2005. Here, the DOJ had sued Dentsply (the dominant manufacturer of prefabricated artificial teeth), claiming that its exclusive dealing arrangements with distributors amounted to illegal monopoly maintenance. The Appeals Court agreed with the DOJ, based on a number of findings. First, relying on Dentsply's stable market share (of around 75–80 per cent based on revenues) over the previous decade, it inferred that Dentsply possessed market power. Second, it established that direct sales to customers (that is, bypassing distributors) did not pose a sufficient competitive constraint on distributors, since the latter were much more effective and entrenched in the system:

[W]e are convinced that direct [selling] is 'viable' only in the sense that it is 'possible', not that it is practical or feasible in the market as it exists and functions [...] That some manufacturers resort to direct sales and are even able to stay in business by selling directly is insufficient proof that direct selling is an effective means of competition. The proper inquiry is not whether direct sales enable a competitor to 'survive' but rather whether direct selling 'poses a real threat' to defendant's monopoly. (Paras 67 and 69 of the judgment)

Therefore Dentsply's exclusivity policies were found to prevent effective competition. Moreover, the Court found additional harm due to more limited consumer choice, since consumers were deprived of competing brands. Finally, the Court ruled that the alleged business justifications were pretextual. We discuss this case in greater detail in Section 3.7.5.

In *McWane*, the Appeals Court agreed with an Opinion by the Federal Trade Commission ('FTC') that a dominant supplier of domestic ductile iron pipe fittings (McWane) unlawfully adopted exclusive terms in its agreements with distributors, in violation of Section 5 of the Federal Trade Commission Act. Specifically, the FTC and the Court found that following entry by a competitor (Star) in 2009, McWane's introduced contract terms

containing restrictive clauses. These stipulated that if any of McWane's distributors wished to purchase fittings from McWane's competitor, that distributor would have forgone the opportunity to purchase any supplies of fittings or accessories from McWane for up to 12 weeks, as well as any rebates that they would have otherwise been entitled to. The FTC and the Court also pointed to internal documents which they interpreted as showing an intent by McWane to contain Star's volumes, so as to prevent it from reaching minimum efficient scale and to make it unprofitable for it to build its own foundry in the US (this would have been more cost-effective in the longer term, according to the FTC), thus ultimately preventing it from becoming a viable competitor.^{71,72} The Court also dismissed McWane's efficiency defences.⁷³

Finally, in *Victrex*, the FTC settled with Victrex's subsidiary Invibio in 2016 in relation to practices concerning the supply of a high-performance polymer used to make spinal and other medical implants. The FTC argued that Invibio used long-term exclusive contracts to prevent its customers from using more than one source of supply for the polymer, implant-grade polyetheretherketone ('PEEK'). The FTC stated that Invibio was the sole supplier of this polymer until two competitors, Solvay and Evonik, entered the market in the late 2000s. The FTC found that Invibio extracted exclusivity terms from customers both by threatening to withhold critical supply or support services and by offering small discounts. It concluded that Invibio's exclusive contracts foreclosed a substantial majority of PEEK

⁷¹ For example, the Court's judgment refers to the following quote by a McWane executive: 'We need to make sure that [Star does not] reach any critical market mass that will allow them to continue to invest and receive a profitable return' (page 7).

⁷² We note, however, that one FTC Commissioner (Joshua Wright) dissented quite strongly with the FTC's majority and – unsuccessfully – requested a review of the Appeal Court's judgment by the Supreme Court. In summary, although Commissioner Wright agreed that McWane's conduct had an exclusionary potential according to the theory of harm put forward, he argued that the FTC did not present any economic evidence to support such a theory. For example, he stated that no estimates of minimum efficient scale of product were presented. To the contrary, he stated that evidence existed (but was not presented) that McWane's exclusivity clauses did not harm competition: in particular, such clauses did not reduce output or increase prices. He also argued that the most plausible inference to draw from the facts of the case was that such clauses had almost no impact on Star's ability to enter and grow its business. He also pointed to evidence showing that Star's growth rate was identical before and after McWane stopped enforcing the exclusivity clauses. See Wright (2015) for further details.

⁷³ Two efficiency defences were put forward for the exclusivity clauses. They were allegedly necessary: to keep McWane's domestic foundry running; and to prevent Star from cherry-picking the core of the domestic fittings business by only manufacturing the top few dozen fittings.

sales from Invibio's rivals. The FTC also noted that there was a significant risk that Invibio's exclusive contracts would preclude Solvay and Evonik from achieving sufficient returns to justify future investments, including in innovative technologies. Under the FTC's final order, Invibio would be generally prohibited from entering into exclusive supply contracts and from preventing current customers from using an alternate source of PEEK, as well as from using other contract provisions – such as market-share or retroactive rebates – that could effectively result in an exclusive arrangement between Invibio and a device-maker.

3.6.2 European Union

To our knowledge, the first time where the Commission and the Courts dealt with exclusive dealing was in *Suiker Unie*. There, the Commission found an abuse of dominance on this count (among others).^{74,75}

In *BPB Industries*, BPB Industries (the dominant supplier of plasterboard in the UK through its subsidiary British Gypsum) offered payments to builders' merchants conditional on exclusivity. The main goal of the scheme according to the Commission was to prevent importers from gaining a foothold in the UK. The Commission deemed such conduct abusive, and the General Court agreed.

A case that diverted from the standard approach in the analysis of exclusive dealing by the Commission and the Courts was *Van den Bergh*. The case concerned HB, the dominant supplier of 'impulse ice cream' in Ireland (these are individually wrapped industrial ice cream products consumed immediately upon a purchase, typically satisfying a sudden desire). HB was found to have a stable market share in excess of 75 per cent and had offered retailers a freezer cabinet free of charge, on the condition that they would only stock HB products there. Roughly 40 per cent of all

⁷⁴ This case was in fact mostly concerned with a large-scale cartel across Europe. However, the Commission also analysed, under Art. 102, the exclusive dealing clauses and fidelity rebates adopted by Südzucker Verkaufsgesellschaft ('SZV'), which it found dominant in the southern part of Germany. The Commission found that SZV's share of the market was at least 90 to 95 per cent and concluded that the exclusive dealing clauses and the fidelity rebates were abusive, without any economic analysis. The Court of Justice, however, annulled the Commission's finding of abuse of dominance based on exclusive dealing clauses, due to the legal form of the relationship between SZV and re-sellers (who were deemed trade representatives, acting as an auxiliary organ of SZV, and not dealers).

⁷⁵ In *Hoffmann-La Roche*, the Commission found that Roche, a dominant producer of vitamins, had required its customers (typically large industrial buyers) to purchase exclusively, or almost exclusively, from it. As this was mainly achieved through the use of rebates conditional on exclusivity across different product categories, we review this case in Chapter 4.

available retailers had signed up to this scheme. HB was found to have infringed both Art. 101 and Art. 102 TFEU (both the General Court and the Court of Justice upheld). As noted among others by O'Donoghue and Padilla (2013: 431), the main novelty from a legal perspective was that on appeal the General Court, after reviewing HB's behaviour under Art. 101, it carried out the Art. 102 analysis in a similar fashion, that is, assessing the likely effects of the conduct in question. On the actual merits of the case, the General Court confirmed that HB's arrangements foreclosed competitors, in spite of consumer demand for alternative brands. This was mostly due to the inability to store a second freezer cabinet in-store: a survey done for the Commission revealed that 87 per cent of the Irish outlets that responded did not find this a viable alternative, because of lack of space. Moreover, 82 per cent of respondents said they would not consider replacing the HB freezer with one supplied by one of HB's competitors (due to the popularity of HB's products). The Commission also ruled out that an outlet with an HB cabinet would have a financial incentive to dispose of it and purchase one itself, stocking the mix of products it wished. Against these market conditions, the foreclosure effects dominated the potential efficiency gains, according to the Commission.

Just over two years after the Court of Justice's judgment in *Van den Bergh*, the European Commission indicated in its Guidance Paper that it would assess exclusive dealing (like other possible abuses of dominance) using a more economics-based approach.⁷⁶ However, in *Tomra*, one may query whether this occurred.⁷⁷ This case was mostly concerned with rebates and is thus reviewed in Chapter 2; however, it also contained some elements of exclusive dealing, which we discuss here. Tomra was a supplier of reverse vending machines, used for the collection and recycling of empty drink containers (it also provided maintenance and repair services to such machines). At the end of the 1990s, it had a market share of high-end vending machines exceeding 95 per cent across the European Economic Area as a whole. Tomra, *inter alia*, was sanctioned for entering into exclusive dealing contracts (varying between one and three years) with customers in Austria, Germany, the Netherlands, Norway and Sweden, from 1998 to 2002. These contracts covered a proportion of total demand over this time period ranging from about 20 per cent to nearly 70 per cent.

⁷⁶ European Commission (2009). See in particular paras 20 and 36.

⁷⁷ Admittedly, the Commission's decision was published in 2006, before the publication of the Guidance Paper referred to. However, the Commission's Directorate-General for Competition had published in 2005 a Discussion Paper setting out similar economic principles – see Directorate-General for Competition (2005).

An important feature of the Commission's decision, and even more so of the judgments by both the General Court and the Court of Justice (both of which were upheld), was therefore the lack of 'a more economic approach' in analysing this case: the exclusive dealing provisions were deemed illegal based on a very formalistic assessment.⁷⁸

In the same vein, it is likely that the *Intel* judgments (see Chapter 2) will also have an impact on the EU policy on exclusive dealing. It seems likely that the tough stance adopted by the judges on 'exclusivity rebates' would *a fortiori* apply to exclusive dealing arrangements.

At the time of writing, there are a number of ongoing investigations by the European Commission involving Google. In the context of practices that may be analysed using the exclusive dealing framework set out in this chapter, there are open proceedings in *Google Search (AdSense)* and in *Google (Android)*. In *Google Search (AdSense)*, the Commission has reached the preliminary view that Google has been abusing a dominant position in online search advertising by artificially restricting the possibility of third party websites to display search advertisements from Google's competitors, thus enabling Google to protect its dominant position. In *Google (Android)*, the Commission has alleged that Google has been offering financial incentives to manufacturers and mobile network operators on the condition that they exclusively pre-install Google Search on their devices, and that this too may amount to an abuse of a dominant position.^{79,80}

3.7 Cases

In this section, we review in some detail a number of cases from a range of jurisdictions. We have selected these because in our view, based on publicly available information, there were certain features that may allow one to

⁷⁸ See, for instance, Federico (2011b), with an emphasis on the analysis of rebates.

⁷⁹ We mention other aspects of the *Google (Android)* case in Chapter 4 in relation to tying. In Chapter 5 (Section 5.6.8), we discuss instead the *Google (comparison shopping)* case.

⁸⁰ At the time of writing, the European Commission is also investigating potentially anti-competitive practices by Qualcomm. In December 2015, the Commission announced that it had taken the preliminary view that, inter alia, Qualcomm's practice of making payments to a major smartphone and tablet manufacturer on condition that it exclusively use Qualcomm baseband chipsets in its smartphones and tablets reduced the manufacturer's incentives to source chipsets from Qualcomm's competitors and harmed competition and innovation in the markets for UMTS and LTE baseband chipsets. See European Commission, press release IP/15/6271 of 8 December 2015 in Case 40.220 *Qualcomm*.

interpret at least some of these cases in the spirit of the economic models presented in this chapter. As such, these cases may provide some helpful illustrations of some of the mechanisms previously discussed.

3.7.1 Langnese-Iglo

This case concerned the ice cream industry. Although it was analysed by the European Commission under what is now Art. 101 TFEU, we find it helpful to review this case in order to discuss the application of some of the economic models set out earlier in this chapter. The facts of the case, based on the information publicly available, were roughly as follows.⁸¹ In 1991 Mars brought a complaint arguing that Langnese-Iglo ('LI') and Schöller had illegally entered into exclusive dealing arrangements with German retailers. The relevant product market was the same as in *Van den Bergh* (discussed in Section 3.6.2), namely impulse ice cream. Clearly, to provide such products to final consumers, retail outlets needed freezer cabinets. However, they typically found them too expensive relative to the revenue stream generated. Industrial ice-cream suppliers, such as LI and Schöller, offered them without charge to retailers, upon the condition that they did not stock competitors' products. LI's exclusivity clauses prevented retailers from stocking competing products in the freezer cabinets provided by LI and more generally in the retailers' outlets. However, the Commission and the Courts failed to make a clear distinction between the likely effects of outlet exclusivity and those of freezer cabinet exclusivity. This distinction is important, as we set out next.

First, consider 'exclusivity' to mean 'outlet exclusivity' (that is, retailers wishing to have a freezer cabinet from LI could not stock any other impulse ice cream in-store). It would be difficult to find an efficiency rationale for such behaviour in this case. A classic pro-competitive argument behind exclusive dealing is that both manufacturer and distributor are genuinely committed to exert effort to promote a certain good or service (thus protecting the initial investment in the product and in building a high-quality distribution network, including, for instance, good staff training); but such defence is unlikely to have a strong impact in the case of packaged ice-creams (differently from, say, advanced technological products). By contrast, the foreclosing effect could in principle be significant. First, scale

⁸¹ Here we draw from Motta (2004: section 6.6.2). The General Court annulled the part of the Commission's decision which prohibited LI from concluding exclusive purchasing agreements in the future. The Court of Justice sided with the General Court. There was a parallel case against Schöller.

economies are likely to be important: there are probably substantial fixed costs in setting up a distribution system and in creating a successful brand (a key factor in this industry). Second, demand was fragmented, with most of the demand (60 per cent) coming from small retailers, thus buying small volumes. Third, contracts were staggered in time, leaving scope for discriminating contracts across retailers. These are all features that make the exclusionary equilibria more likely, as we saw in Section 3.2.3. On the actual merits of the case, LI's exclusivity agreements covered something in excess of 15 per cent of the relevant market; together with Schöller, Motta (2004: Section 6.6.2) noted, the agreements between these two players foreclosed about 30 per cent of the relevant market, as defined by the Commission, with an average contract duration of two and a half years.⁸² Hence, there was a lot of contestable demand, potentially making a theory of harm based on exclusive dealing less cogent; yet we are not told anything in the Commission's decision as to whether this 30 per cent of uncontested demand included key distributors in the market; in other words, could (efficient) ice cream suppliers be viable in the business simply by addressing the contestable demand? The answer seems to be positive, given that LI and Schöller faced several competitors with non-negligible market shares.

By contrast, consider 'exclusivity' to mean 'freezer cabinet exclusivity'. The efficiencies from such contracts would be sizeable: small retailers would not purchase one without financial support; and industrial ice cream suppliers would not offer a cabinet free of charge if competitors could free-ride on their investment and place their products there too (this would be an example of the investment promotion having a 'positive external effect' – see the discussion in Sections 3.3.1, 3.4.6 and 3.5 – thus depressing the incentive to invest absent exclusivity). Foreclosure effects, on the other hand, would be likely to be limited: after all, even in a small store, there was the possibility of stacking an extra small freezer (for example vertically), and Mars adopted these after the Commission's decision (the Commission was instead more sceptical in relation to in-store space availability at the time).

In sum, even abstracting from the somewhat dubious market definition of this case and the lack of detailed evidence one may draw the following conclusion: if the main concern was about outlet exclusivity, there may have been some evidence consistent with a theory of harm based on scale

⁸² One possibly controversial issue with this case was market definition: the Commission considered that craft-made (or scooping) ice cream should not be part of the relevant market, but we do not dwell on this point.

economies (although there was also some evidence inconsistent with it, as we discussed); whereas if the main concern was on freezer cabinet exclusivity, it would be very difficult to find much evidence to support a robust theory of harm.⁸³

3.7.2 TIM

In *Telecom Italia – Costituzione rete dealer GSM* (or *TIM*), the Italian competition authority ('AGCM') ruled against Telecom Italia Mobile ('TIM'), the incumbent mobile network operator in Italy. The case referred to conduct in the mid-1990s, that is, at the outset of the commercial era of GSM (or second-generation) mobile telephony services. At the time, according to the AGCM, direct distribution (that is, directly from the mobile operator) was only used to sell GSM services to large corporations and institutional clients. Retail consumers were instead reached via independent dealers. As for TIM in particular, it had signed exclusive dealing arrangements (of up to three years) with 1,875 dealers across the country; such indirect distribution accounted for 85 per cent of the new GSM subscribers between July and October 1995. It is also useful to recall that at the time, in Italy, there was hardly any pre-pay, or pay-as-you-go, mobile service nor any SIM-free handset; that is, an end-user wishing to use such services essentially had to buy a handset together with a contract with a mobile network operator. This was also due, as noted by the AGCM, to large handset subsidies offered to new subscribers.

TIM was found to have a market share in excess of 80 per cent (by number of subscribers). The AGCM was concerned that TIM's exclusive contracts with the dealers would foreclose nascent upstream competition. Omnitel, the only competitor at the time, had just started its operations and retailed its services both directly and through independent dealers, though without exclusive agreements. However, the AGCM noted, there was a significant asymmetry between dealers in terms of retailing potential. Out of around 10,000 mobile phone dealers, 1,940 signed 78 per cent of all mobile services contracts; the success of such dealers was due to a mix of commercial skills, attractive product lines and favourable location. The AGCM pointed out that an 'enormous majority' of these 1,940 outlets coincided with the 1,875 dealers with which TIM had signed exclusive

⁸³ Note, though, that as the Commission believed that there were limited possibilities for a second in-store freezer cabinet, 'freezer cabinet exclusivity' and 'outlet exclusivity' would be likely to coincide under that belief, at least in the short term.

agreements. By contrast, the AGCM continued, Omnitel was left with access to less specialised retailers, who were less effective (on average they each signed 52 subscriptions to end-users over the four months leading to the inquiry, as opposed to TIM's 178). Omnitel was therefore forced to raise the number of dealers (it had 2,100) to counteract a lower average dealer 'quality'; this was not negligible in an industry where distribution costs were estimated (at least in the US) to hover at around 35 per cent of total costs (Omnitel argued that its distribution costs were 54 per cent higher than what would have occurred absent TIM's exclusionary behaviour). The AGCM was also concerned by the financial incentives provided to dealers in return for exclusivity. Most specifically, it referred to the rebate given to the (exclusive) dealers, amounting to 1 per cent of the revenues from traffic of all end-users signed up by that dealer (provided a minimum of 125 subscriptions were signed). The AGCM also rejected a possible efficiency defence in the form of investment protection (that is, the need to avoid free-riding, by competitors, of know-how passed by TIM to its dealers): it found no evidence that TIM used to train its exclusive dealers or undertook any significant relationship-specific investment.

Against this background, one could argue that some of the facts presented by the AGCM in this case are consistent with the factors that may underpin a theory of harm based on scale economies, as presented in Sections 3.2.3 and 3.4.3. First, economies of scale played an important role. The fixed costs of setting up and running a mobile phone network, especially at its infancy, are very high, and likewise advertising costs. Second, TIM had a significant incumbency advantage, as noted by the AGCM, which also referred to TIM having been part of the same group owning the incumbent fixed telephony network and having used over 1,000 of the same dealers as TIM's to retail fixed telephony products. Third, buyers (dealers/retailers) were very numerous and clearly fragmented.

These conditions are therefore potentially consistent with those presented in technical Section 3.4.3.1: while certain dealers were more important than others (because they sold more), no single dealer was sufficient to support entry, or significant expansion, by an upstream competitor (a mobile network operator). So, while TIM had an incentive to mainly focus on these more important dealers, it arguably did not need to offer sizeable compensations to involve those buyers in exclusivity (1 per cent of revenues from the mobile network traffic of the customers signed up), as the model set out in Section 3.4.3.1 would predict.

3.7.3 Posten Norge

Posten Norge is a case of exclusive dealing by a near-monopolist. In July 2010 the European Free Trade Association Surveillance Authority (the ‘Authority’) ruled that Posten Norge, the Norwegian incumbent postal operator, had abused its dominant position in the market for over-the-counter delivery of parcels sold by distance-selling businesses to Norwegian consumers between 2000 and 2006. The specific conduct that was found to be abusive consisted of the exclusive agreements that Posten Norge signed with selected retailers (for example, supermarket chains) which would provide Posten Norge with on-store space to offer over-the-counter postal and financial services on an exclusive basis.⁸⁴ We next summarise the facts of the case, as set out in the Authority’s decision, before commenting on them.

Privpak, the complainant, was a Swedish company offering delivery services to distance-selling businesses, by carrying their parcels to consumers in Norway, Sweden and Finland (B-to-C parcel services). Its business model relied on transporting parcels to retail outlets (for example, grocery stores and petrol stations, for later collection), as opposed to home deliveries. In Norway, such over-the-counter deliveries were much more popular than home deliveries (in the case of Posten Norge, by a ratio of about 24 to 1). This business model relied on the availability of a granular and extended network of existing available outlets. Downstream entry (that is, setting up a dense network of outlets just to deliver B-to-C parcel services) would have been very costly.

Posten Norge had started a cost-reducing strategy of replacing its Post Offices with a network of Post i Butikk, or Post-in-Shops (that is, small branded corners within a retailer’s premises offering Posten Norge’s postal and financial services). According to the Authority’s decision, in 1997 Posten Norge had 267 Post-in-Shops and 1,269 Post Offices; in 2006, these figures were almost reversed (that is 1,184 Post-in-Shops and 327 Post Offices). In establishing this new network, Posten Norge signed exclusive agreements with the three largest food retailers in Norway (NorgesGruppen, ICA and COOP), the main small grocery store chain (Mix, owned by NorgesGruppen) and one of the leading petrol stations chains (Shell). Over the period of the infringement, Posten

⁸⁴ The European Free Trade Association Court upheld the finding of an abuse of a dominant position.

Norge had over 97 per cent of the Norwegian market for B-to-C parcel services.

As for the potential efficiencies arising from exclusivity with the retail chains, the Authority considered some potential objective justifications: exclusivity was a possible solution to the hold-up problem (where the upstream supplier needs to make an initial investment) as well as to the free-riding problem (where distributors do not contribute towards the initial cost of building a brand or some infrastructure, or towards the training of staff, who may then use the acquired skills when supplying competing products). The Authority discounted these efficiencies insofar as group exclusivity was concerned: Posten Norge did not provide evidence as to why the whole NorgesGruppen, for instance, had to sign to exclusivity for such efficiencies to arise. As for outlet exclusivity, the Authority continued, the actual quantitative effect of such efficiencies was likely to be small and would in any case be outweighed by the anti-competitive (foreclosing) effect due to exclusivity.

Against this background, one may find a number of features that could make one interpret this case in the spirit of the main theory we proposed in this chapter (see Sections 3.2.3 and 3.4.3).

First, Posten Norge had a sizeable incumbency advantage through its postal delivery operations. Further, it may have gained an additional first-mover advantage when it began switching from Post Offices to Post-in-Shops back in 1997; this was almost exactly at the time when Privpak began considering entering the Norwegian market – though its operations only started in 2001. This timing is interesting: Posten Norge began negotiations with retailers (including exclusivity agreements) between 2000 and 2001.⁸⁵

Second, scale economies were a key industry factor. On the cost side, the Authority noted:

The [...] market is also characterised by significant economies of scale [...]. Norway Post has realised scale economies in the picking up, sorting and transport of B-to-C parcels. In addition to the scale economies in the transport between terminals, there are scale economies in the transportation of B-to-C parcels from terminals to delivery outlets. (Para. 417 of the decision)

⁸⁵ Posten Norge actually approached other retailers than those mentioned earlier, but it reached no agreement: one retailer was a discount store wishing not to increase its costs, another was a small grocery chain believing that Post-in-Shops would not increase customer flows and both Hydro Texaco and Esso, two petrol chains, argued they did not have sufficient space to accommodate Post-in-Shops.

Scale economies were also very prominent on the demand side, the Authority noted, since this was a multi-sided market. Posten Norge and Privpak were ‘platforms’ trying to attract three types of agents: distance-sellers (for example, e-commerce businesses), final consumers (who could sometimes choose between platforms on a given purchase) and retail outlets. The more agents on a given side joined a platform, the more this was valuable to the other two sides: for instance, an e-commerce selling classic music records would have much less of an incentive to sign a distribution contract with Privpak if it knew that this platform was only available at a handful of retail locations across Norway; many consumers would have probably been unwilling to purchase music records from such e-commerce simply because there was no collection point nearby. The remaining retail outlets (that is, those not covered by the exclusivity agreement with Posten Norge) may have been either insufficient for Privpak to gain a critical scale or, in any case, they may have not been interested in this business opportunity. Moreover, distance-selling companies regarded Privpak’s network as limited (though its prices were found to be competitive).

As noted above, Posten Norge secured exclusivity with the three largest retail groups, in addition to the main small grocery stores chain and a major fuel retailer. This severely limited the ability for Privpak or other competitors to reach a critical mass of retail outlets.

In the context of the models discussed in technical Section 3.4.3, a key question would be whether a single retail chain would have been sufficient to sponsor entry (or guarantee viability to Privpak). Each of the three largest chains roughly had the number of outlets that Privpak estimated would be required for it to be competitive and reach about 90 per cent coverage.⁸⁶ Therefore, in the context of the model with asymmetric buyers discussed in technical Section 3.4.3.2, the incumbent may have potentially found it sufficient, in order to exclude, to offer large compensations to such critical buyers. It is unclear based on information in the public domain what the exact level of compensation was for the retailers. However, the

⁸⁶ The Authority used a working definition of ‘coverage’ as proportion of households that an over-the-counter delivery network could reach (see footnote 29 of the decision). See also paras 234–5 of the decision reporting Privpak’s estimates. Of course, to achieve a certain level of coverage, not only the number but also the *locations* of a retail chain’s outlets would matter. So without detailed information, one cannot conclude that a retail chain with 1,067 outlets (the number estimated by Privpak to reach 89 per cent coverage) would actually guarantee 89 per cent coverage.

Authority clarified that Posten Norge remunerated individual outlets for every postal or financial transaction made on-premise (that is, not just for B-to-C parcel services), in addition to the payment of a fixed monthly fee to cover training, insurance and accounting. There was also an overall pre-agreed minimum monthly payment that Posten Norge guaranteed to the outlets. Moreover, it is also unclear whether the incumbent could recoup the potential losses suffered on the critical buyers through the margins obtained from the remaining buyers. However, based on the Authority's decision, it also appears that notwithstanding the financial remuneration from Posten Norge, retailers were particularly keen to deal with Posten Norge thanks to the increased footfall (and revenues) that it brought to their stores. In fact, retailers actively sought to achieve 'preferred partner status' with Posten Norge (for example, with a guarantee of some retailer exclusivity or priority in certain local areas). In a sense, therefore, a preferred partnership was used as part of the compensation offered by Posten Norge, and may have allowed Posten Norge to have exclusivity accepted at relatively low cost. This aspect also raises the question as to whether exclusive dealing also ought to be assessed, occasionally, as an agreement between firms, with the potential to restrict competition, as we noted in Section 3.5.

3.7.4 Lorain Journal

This is an old case from the 1950s, brought by the US Department of Justice. The defendant (the Lorain Journal Company) had refused advertisements from firms that also advertised through a local radio station ('WEOL') in the area of Lorain, Ohio.⁸⁷

Lorain Journal was found to have a 'substantial monopoly in Lorain of the mass dissemination of news and advertising, both of a local and national character' (para. 147 of the judgment), from 1933 to 1948. It reached 99 per cent of the households in the local area. WEOL began operating locally in 1948. As noted by the Court, 'WEOL's greatest potential source of income was local advertising. Loss of that was a major threat to its existence' (para. 153). Many local advertisers would have liked to advertise through both media. However, confronted with *Lorain Journal's* policy, they could not afford discontinuing print advertisements, given the ubiquitous reach of the newspaper in the local area. Based on such market conditions, the Supreme Court found a breach of Section 2 of the Sherman

⁸⁷ See Gavil et al. (2008: 588 *et seq.*) for a more thorough discussion of the case.

Act. Carlton (2001) is fairly sympathetic with that finding: if scale effects at early stages (determining the growth rate of the nascent competitor) were substantial, then the theory of harm proposed in that case was coherent.

This case can also be read in the spirit of the model we proposed in Sections 3.2.3.1 and 3.4.3.1. First, *Lorain Journal* may have had a significant incumbency advantage given by its publication as a key outlet for advertisers in the local market. Second, scale economies on the demand side were very significant: the media are classic two-sided markets; the local radio station, to be viable, needed a critical mass of advertisers and listeners. By thwarting the presence of the first advertisers, the development of the whole platform was potentially undermined (this is essentially the point made by Carlton (2001), which we referred to above). Third, the customers (local advertisers) appeared to be fragmented and dispersed, so a ‘divide-and-conquer’ strategy was feasible.

3.7.5 Dentsply

We briefly described the facts of *Dentsply* in Section 3.6.1. Here, we delve deeper into the facts of the case and try to interpret it in an economically coherent way.⁸⁸

Dentsply was the dominant manufacturer of prefabricated artificial teeth. It sold these products to dealers (or distributors) who would in turn supply the teeth (and other materials) to dental laboratories (which would then fabricate dentures for sale to dentists). According to the records, there were many dental dealers across the US, some of local, others of national scope; and about 7,000 laboratories capable of fabricating dentures. The Appeals Court, based on Dentsply’s stable market share (of around 75–80 per cent by revenue) over the previous decade, found that Dentsply had market power. Foreign manufacturers were not found to pose a strong competitive constraint, as they failed to adapt their products to American consumer preferences.

At issue were Dentsply’s exclusive contracts with dealers (which were terminable at will). The crucial point of contention in the proceedings before both courts was whether direct selling (from manufacturers to laboratories) provided a competitive constraint on a dealer network (that is, that of Dentsply). The District Court had held that this was the case, based on oral evidence by industry witnesses stating that avoiding

⁸⁸ On some of the facts of the case, we borrow from Sher and Russell (2005), as well as from the judgment of the District Court.

intermediaries would cut costs. The Appeals Court, by contrast, ruled that direct sales did not pose a sufficient competitive constraint on dealers, since they in fact generated efficiencies and provided additional services. In particular, it highlighted many value-added, cost-reducing services offered by dealers.⁸⁹

Let us move to an analysis of the case in the spirit of the main model of this chapter, outlined in Section 3.2.3. The first element to point out is Dentsply's incumbency advantage, having been the dominant market player for over 10 years. Its full line of products, strong brand and reputation made its offer very important for a dealer.

As for the other crucial ingredients of the model that we presented in Section 3.2.3 (and in Section 3.4.3), there was no information in the judgments that would allow us to form a view on whether scale economies were sizeable; moreover, buyers' fragmentation did not seem to be strong: Dentsply's top two dealers accounted for close to two-thirds of Dentsply's total sales (and therefore about half of the overall market). It is then conceivable that one of the largest two dealers may have been sufficient to allow for entry or significant expansion by one of Dentsply's competitors. We do not have sufficient information on whether also the remaining dealers were large (thereby making exclusion unprofitable) or whether there were significant asymmetries among dealers such that Dentsply could have offered discriminatory terms to large dealers and could have earned sufficient margins on the remaining ones. However, from the facts of the case, it appears that it was not so much 'compensation' that mattered, but the consequences from not adhering to exclusivity.

Indeed, one element that may raise concerns in the *Dentsply* case was the existence of an 'all-or-nothing' clause (formalised in the so-called

⁸⁹ '[Laboratories] buy far more heavily from dealers than manufacturers. This may be largely attributed to the beneficial services, credit function, economies of scale and convenience that dealers provide to laboratories, benefits which are otherwise unavailable to them when they buy direct. [...] [T]hey provide laboratories the benefit of one stop-shopping and extensive credit services. Because dealers typically carry the products of multiple manufacturers, a laboratory can order, with a single phone call to a dealer, products from multiple sources. [...] Buying through dealers also enables laboratories to take advantage of obtaining discounts. [...] Another service dealers perform is taking back tooth returns. [...] [U]sing dealers, rather than manufacturers, enables laboratories to consolidate their returns. In a single shipment to a dealer, a laboratory can return the products of a number of manufacturers, and so economize on shipping, time, and transaction costs. [...] Dealers also provide benefits to manufacturers, perhaps the most obvious of which is efficiency of scale. Using select high-volume dealers, as opposed to directly selling to hundreds if not thousands of laboratories, greatly reduces the manufacturer's distribution costs and credit risks' (paras 58–64 of the Appeal Court's judgment).

Dealer Criterion 6 of Dentsply's policy). Dealers adding competitors' teeth to their product line would lose the entire set of Dentsply's products, which included a wide set of artificial teeth, some of which represented 'must-carry' products, and other dental products. As discussed in Section 3.5, if smaller rivals are unable to contest a significant proportion of the demand from buyers because, as it appears in this case, they do not provide the full line of products or because they are capacity-constrained, then an 'all-or-nothing' clause makes the buyers' value of rejecting exclusivity extremely low. This makes it much easier for the dominant firm to induce buyers to accept the exclusive contract.

We have already pointed out in Section 3.5 that, in the short-run, the threat not to supply the buyer at all if it violates exclusivity is not credible. If exclusivity has been violated, and some of the buyer's requirements are left unsatisfied by the rival, why should Dentsply deny to supply its products? However, the records of the case contain several examples of situations in previous periods where Dentsply actually carried out such threats. Hence, it seems that Dentsply had been capable of building a reputation for enforcing an 'all-or-nothing' clause. This may have allowed Dentsply to be successful in imposing exclusive contracts to dealers and to exclude rivals even if such contracts had short duration (given they were terminable at will).

3.7.6 Nielsen

Another major case from North America concerning exclusive dealing is *Nielsen*, on which the Canadian Competition Tribunal ruled in the mid-1990s. D&B Companies of Canada (or Nielsen, after the name of the relevant affiliate company) was found to be the monopoly supplier of scanner-based market tracking services in Canada. These services involved the collection of data on retail sales over time, at a highly disaggregated level (purchase-by-purchase). The data was then analysed, aggregated as appropriate and elaborated by a specialised firm (such as Nielsen). This firm finally customised reports for manufacturers (for example, of food), enabling them to optimise their marketing based on recently gathered information. (The defendant heavily disputed this market definition, arguing that other marketing services, such as audits and consumer panels for instance, were also part of the same relevant market.)

This case is very interesting as the dominant firm was active at the intermediate level of a vertical chain and it engaged into exclusive contracts both upstream and downstream. Its upstream relationship was with

retailers (for example, supermarkets), from which it purchased data. Its downstream relationship was with manufacturers, to which it provided marketing services based on the elaborated data.

Downstream relationships Let us focus on the *downstream* relationship first. Nielsen's incumbency advantage in the mid-1990s was well established in Canada according to the records. It had been the sole supplier of scanner-based market tracking services (or similar services with a less sophisticated technology) for a decade. This incumbency advantage was compounded by a first-mover advantage. As Jing and Winter (2014) noted, the terms of Nielsen's contracts with manufacturers increased from less than one year (evergreen contracts with eight months' termination notice) to three to five years at the time when IRI sought to enter the Canadian market. These contracts also contained liquidated damages payable to Nielsen if a manufacturer terminated the contract. The Tribunal found that the purpose of such contract length was to 'lock-up' customers in long-term contracts (this phrase was also found in internal business documents).

As for the scale economies in this case, the Tribunal assessed what elements a potential competitor would need in order to enter the market, namely: technology and know-how required to put together a representative sample and to collect, process and analyse large volumes of data; a Canadian-specific dictionary that translates the product code into detailed descriptions useful to tracking service users; a field-force that collects the data from the stores; the actual availability of scanner data from retailers; and the availability of potential customers.

Another main ingredient was the presence of some important (downstream) buyers: US manufacturers retailing in Canada which purchased services from IRI, a competitor of Nielsen in the US. From the records, it emerged that Nielsen had a strong interest in engaging in exclusive contracts with such buyers (in Canada), as these would have otherwise facilitated entry into Canada by IRI.

Taken together, these factors were likely to give rise to contracting externalities between buyers, the essential element of the model discussed in Sections 3.2.3 and 3.4.3, where a single buyer would be insufficient to support upstream entry (a notion we have been referring to as 'buyer fragmentation'). As noted by Professor Ralph Winter in his Expert Witness Statement, manufacturers had a collective incentive to sign short-term contracts with Nielsen, to reduce their exposure to its market power.

However, individually, they each had an incentive to accept long-term exclusivity in return for a discount; in accepting such an offer, each buyer was not internalising the effect on the other buyers, exposing them to stronger market power by Nielsen.

Upstream relationships Let us next turn to the relationship between Nielsen and input suppliers, that is, the *upstream* exclusive contracts with retailers. The exclusionary mechanism just described for the downstream relationship may have actually been reinforced by the upstream (input) foreclosure that occurred: by signing all major grocery retailers into exclusive agreements since 1986, Nielsen reduced the ability of a potential entrant into the supply of scanner-based market tracking services to obtain data from retailers, which was clearly an essential input in that business. This would in turn weaken the potential entrant's position in the market vis-à-vis the non-exclusive customers.

The Tribunal insisted that Nielsen's exclusionary strategy was particularly effective thanks to the use of contract staggering at the upstream level:

Nielsen's proven strategy of staggering contract renewals further reinforces that the purpose of the exclusive retailer contracts was and is to exclude potential competitors. [...] Nielsen was concerned with staggering the contracts in order to minimize the payments to retailers. (Page 65 of the judgment)

The effect of contract staggering may have been particularly significant for potential competition, as it occurred both upstream and downstream. The following quote (which is an extract from a Memorandum by a Nielsen executive, reported in Winter's Expert Witness Statement) explains such mechanism:

1. After we did our retail deals five years ago, we recognized that we were vulnerable because virtually all of these agreements expired around the same time. We set ourselves a goal then to pursue a practice that would result in our retailer and distributor contracts expiring at different times. This would make it much more difficult for any competitor to set up a service unless he was prepared to invest in significant payments before he had a revenue stream.

2. Late last Fall we executed a couple of important renewals which, frankly, made it impossible for anyone else to produce a national tracking product for the next five years. (Page 55)

The case also featured 'key' upstream suppliers; the Tribunal pointed out how Nielsen itself was aware of their presence:

It is clear from the documentary evidence that Nielsen was of the view that no firm could produce a national scanner-based market tracking service without access to Canada Safeway's scanner data. (Page 68)

This last point is akin – in the context of the model we presented – to the incumbent preventing the entrant from achieving minimum viable scale in the relevant market.

The theory of harm in this case was therefore fairly elaborate. In fact, one could not argue that Nielsen enjoyed a first-mover advantage when dealing with upstream suppliers. Winter, in its Expert Witness Statement supporting the Government's case against Nielsen, noted that Nielsen, back in 1986, engaged in a bidding war with IRI for exclusivity for large retailers' data (competition *for* the market). One way of interpreting the facts of the case is that *once* the upstream market had tipped in Nielsen's favour, then the standard 'naked exclusion' mechanism (introduced in Section 3.2.3.1) arose. As for the reason why IRI was excluded at the industry's infancy, when both Nielsen and IRI were competing for retailers, one can refer to Jing and Winter (2014), that we mentioned in Section 3.2.5. Upstream inputs were complementary, meaning that the value of a dataset containing information from five large retailer chains, say, was larger than the sum of the values of the five separate databases, each containing information from a single retailer. This leads to the fact that a market outcome where some retailers supply Nielsen while others supply IRI may not be possible because it would determine a relevant efficiency loss. Further, downstream competition between IRI and Nielsen would transfer to buyers (that is, manufacturers) most of the benefits generated when all the retailers supply both IRI and Nielsen. Then, it follows that it would have been in the joint interest of the contracting agents (IRI, Nielsen, and the retailers) that either of the two firms was excluded. Which of the two firms wins this upstream competition depends on the benefits that can later be realised, which is also a function of downstream competition. In the *Nielsen* case, Jing and Winter (2014) find that Nielsen's downstream contracts gave it a stronger incentive than IRI to win the upstream competition, which therefore tipped in favour of Nielsen.

As for the actual outcome of the case, the Canadian Competition Tribunal prohibited Nielsen from entering into exclusive contracts with retailers; it also ordered that all existing contracts with customers (that is, manufacturers) be terminable on eight months' notice.

3.7.7 Distrigaz

Following the European Commission's investigation of the energy sector,⁹⁰ a large number of important decisions were taken. We review some of these in Chapter 5 in the context of vertical foreclosure. Here, we look instead at one commitment decision in the context of exclusive dealing.⁹¹ As the decision is fairly brief, we are unable to reach a definite conclusion as to the validity of the theory of harm on the facts of the case; however, we will suggest that the market conditions were potentially conducive to the exclusionary mechanism we described in Sections 3.2.3.1 and 3.4.3.1.

In 2007, the Commission adopted a commitment decision with respect to Distrigaz, the former Belgian incumbent gas supplier (at the time it was being divested as part of the Gaz de France merger with Suez). Under scrutiny were Distrigaz's long-term contracts, in the market for (high-calorific) gas to large industrial users.⁹² The incumbent (together with an affiliate) had about three-quarters of the relevant market. Barriers to entry were deemed high.⁹³ Distrigaz's contracts with large buyers had either an explicit exclusivity requirement or minimum volume purchase commitments. The Commission provided a snapshot of the proportion of demand tied to Distrigaz, taking 1 January 2005 as the starting point: over half for the following six months, about 40 per cent over 2005 as a whole and about one-quarter over a three-year period. It is debatable whether this amounted to substantial foreclosure, in the sense that a competing gas supplier could not (vially) serve the contestable share of the market. As for scale economies, the barriers to entry referred to earlier, together with the upstream natural monopoly conditions, are potentially conducive to sizeable scale economies. The Commission also summarised its concerns with regards the viability of competitors as a function of the available demand:

⁹⁰ See European Commission, Press Release IP/05/716, 13 June 2005.

⁹¹ Under European Council Regulation 1/2003, Article 9, parties may offer commitments (sometimes also known as undertakings), and if these address the competition concerns by the European Commission, the Commission may close its investigation without an infringement finding.

⁹² Distrigaz's customers included gas re-sellers, that is, competitors; as this would entail potential vertical foreclosure, which we deal with in Chapter 5, we do not dwell on this aspect of this case in what follows.

⁹³ The Commission listed the following barriers: 'the balancing regime on the transport network; the difficulty in using gas in transit to supply customers in Belgium; congestion on the entry points into the Belgian gas transport network; the lack of liquidity on the Zeebrugge Hub which made it an unreliable source of gas for suppliers in Belgium; and the lack of effective competition in the L-gas market, which spilled over to the H-gas market [...]' (para. 15).

[A]ccess to customers could be foreclosed due to the combination of two factors: the duration of the contracts and the volumes of gas tied to Distriga[z]. Alternative suppliers could therefore find it difficult to build up a viable customer base. (Para. 18 of the decision)

The decision does not give details either as to the extent of demand fragmentation. While the main buyers concerned were large industrial customers, Distrigaz nevertheless sold to different industries; it is therefore reasonable to expect that there was a substantial level of fragmentation in the demand faced by Distrigaz. The market also clearly featured staggered demand. The decision stressed that (bar very few exceptions) industrial customers dealt with a single gas provider at a time. That meant that there was a clear date when competition for a client would arise; but such dates were staggered, that is, they differed from one customer to another, also because Distrigaz itself offered contracts of different duration. This may have enabled Distrigaz to make discriminating and sequential offers, or in any case to exploit the market sequentiality, in the fashion we described in Section 3.2.3.

Finally, we cannot comment upon the level of downstream competition: industrial customers operated in different sectors and we are not aware of the level of competition in each of them, nor of the proportion of total costs accounted for by gas as a production input. The decision makes no reference either to the difficulty of renegotiating the contracts, that is, if there were substantial transaction costs.

The key undertakings offered by Distrigaz (and accepted by the Commission) were the following: contracts would be such that a minimum of 65 per cent of the gas supplied to industrial customers would return to the market every calendar year (that is, this was set as the minimum contestable share of total market demand); no single contract would exceed a five-year duration; supply agreements with resellers could not exceed a two-year duration nor include resale restrictions. Importantly, recognising the pro-competitive effects of exclusive dealing in the presence of uncertain investments and hold-up problems, the Commission carved out an exception in the undertakings, insofar as agreements for the supply of gas for new investment in large electricity generation capacity were concerned.⁹⁴

⁹⁴ The facts in *EDF* (another commitment decision concerning the French electricity incumbent) were very similar to those just discussed in *Dstrigaz*, so we do not discuss it further.

3.7.8 Events Ticketing

An industry which has received some antitrust attention is the outsourced ticketing services one. Firms active in this industry provide a platform where consumers can buy tickets for major events (for example, concerts, theatre plays, musicals, sport games); bookings can typically be made online, via the phone or via ticketing outlets (often located within music stores). In what follows, we summarise and comment on the conclusions drawn by the Irish and the Singaporean competition authorities following their reviews of exclusive dealing practices by the dominant providers of outsourced ticketing services.

Let us begin though with a brief description of the basic industry mechanics from an economic perspective (based on the information available in the authorities' decisions). To fix ideas, think of an internationally known rock band. When the band goes on a tour, promoters across countries compete to attract the band to include their country in the tour. Promoters are responsible to arrange venues and marketing services, as well as ticketing services. Such services are therefore economic inputs for a promoter. The rock-band and the promoter jointly set the face value of the ticket, which typically include details on the artists' fees, as well as the venue costs and marketing.

Competition authorities (see below) have identified a number of entry barriers in the outsourced ticketing services market, as several facilities are needed: a granular network of retail outlets; a call-centre and online ticket sales operations able to deal with high demand at a given point in time; and high-quality hardware and software for ticketing.

In the Irish case, the Irish Competition Authority (the 'Authority') did not find any breach of the competition rules, in spite of TicketMaster being a monopoly provider of outsourced ticketing services for events of national and international appeal on the island of Ireland.⁹⁵ The key reason of this finding was that the two largest promoters in the market (MCD and Aiken) exerted very strong buyer power vis-à-vis TicketMaster in Ireland. As these promoters were competing against promoters in other countries to attract artists, they had a strong incentive to minimise the booking fees charged by TicketMaster, so as to be able to offer higher fees to the artists (for

⁹⁵ To be more precise, the Authority dismissed the excessive pricing complaints on the basis that TicketMaster did not possess market power due to the buyer power it faced; and it found that TicketMaster's contracts with the promoters did not constitute anti-competitive agreements.

a given final price to the consumer). Their agreements placed a cap on TicketMaster's booking fees.

The Authority also found that the two promoters could credibly switch to alternative providers. First, they could outsource ticketing services to an existing retail network (for example, post offices). Second, they could approach existing ticketing services providers of regional appeal. Third, they could sponsor new entry, by guaranteeing viability to an entrant (the two promoters jointly accounted for between 50 and 100 per cent of the entire market). Fourth, they could also self-supply in principle. Further, the Authority believed in the presence of substantial efficiencies arising from exclusive contracts with a single ticketing service provider: lower transaction costs (mostly by not having to tender services for each event); economies of density of a granular outlet network; risk-sharing efficiencies due to upfront and advanced payments that TicketMaster made to the two main promoters;⁹⁶ and lower search costs for consumers.

Let us make a few remarks on the decision. First of all, it appears that the exclusive contracts had been proposed by the two main promoters, which also often called for renegotiation of their terms, and derived considerable benefits from them. This is therefore very different from a situation where it is the supplier that imposes an exclusive contract. Also, some of the above-mentioned efficiencies are simply due to the buyer preferring to deal with just one supplier.

Second, these contracts may well be consistent with a situation – highlighted in several points in this chapter – where large buyers (in this case, the two main promoters) and the dominant incumbent (here, TicketMaster) share the gains from excluding other ticket service providers. Or, more in line with the Authority's reasoning, they may be incompatible with anti-competitive exclusion, if the rents that can be extracted from the remaining promoters are insufficient to compensate the two large ones. Given that the decision just says that the two main promoters accounted for between 50 and 100 per cent of the market, it is difficult for us to discriminate between these two hypotheses.

Third, the Authority seems to argue that this is a market where it is optimal to have just one supplier, due to the importance of having one very granular network, which would also minimise consumers' search costs

⁹⁶ These payments could help raise the large sums of money needed to finance major events; the Authority argued that it was probably more efficient for promoters to raise finance through a company like TicketMaster than through other channels, such as banks, which would have typically been less knowledgeable about the business.

(the decision says that if there were several suppliers, points of sale would be more fragmented and therefore might not be profitable, which in turn might make it harder for some consumers to purchase tickets; see para. 2.89). Apart from noting that there are countries in which several suppliers of outsourced ticketing services co-exist without apparent inefficiencies, there seems to be a contradiction with the Authority's arguments that new entry sponsored by promoters could easily occur, whether through self-supply, by providers of ticket services for regional events, or by international firms.

The Singaporean Competition Commission (the 'Commission') took a different perspective and concluded that SISTIC, the leading ticketing services provider in Singapore (with a market share in excess of 90 per cent), abused its dominant position, because of its exclusive dealing agreements with two major venues and with 19 event promoters (including the two venue operators, which were also promoters).⁹⁷ One of these two venues (the Esplanade) was held by TECL, which itself part-owned SISTIC; the Esplanade was considered by event promoters to be the unique premier location in Singapore, due to its location and excellent facilities and equipment. The other, SIS, was the largest indoor venue in Singapore and attracted about three-quarters of all concert and sport games taking place in Singapore. Alternative ticketing services providers were not believed to impose a competitive constraint. From a survey of event promoters it emerged that such competitors lacked a substantial network size and strong brand. Self-supply (of ticketing services) by event venue operators or promoters was not considered to be viable either.

The Commission conceptualised the industry in a different way from what we presented above: it noted that a ticketing services provider was the 'middleman' between event promoters and ticket buyers, selling services to both sides. It thus portrayed the market as being two-sided (where the price *structure* mattered as much as the price *level*). The Commission emphasised that SISTIC exerted its market power by devising the profit-maximising price structure, instead of being a price-structure-taker. However, the Commission also noted that barriers to entry and scale economies were 'moderately high' (but no more) in this industry, with the exception of indirect network effects, which were reinforced by SISTIC's exclusive agreements. So removing those agreements could make the market more competitive, the Commission concluded.

⁹⁷ The Competition Appeal Board of Singapore upheld.

One important difference with the Irish case (which the Singaporean Competition Commission itself highlighted) was also buyers' fragmentation: in Singapore, no single event promoter had a share of sales in excess of 5–15 per cent; SISTIC had exclusive arrangements with 19 of them but actually dealt with about 200 event promoters overall. As we saw in Sections 3.2.3.1 and 3.4.3.1, buyer fragmentation is a key ingredient of the exclusionary mechanism we presented there. In the Irish case, there was no such buyer fragmentation and in fact the significant buyer power resulted in an alternative platform being seen as credible/viable by the Irish Competition Authority. In Singapore, this was not possible; in fact, SISTIC was found to have exploited the buyer fragmentation to select which partners should be exclusive and then made them discriminatory offers (see para. 7.7.1 of the decision) and staggered (para. 7.9.10). The Commission itself referred (Appendix 7.1) to the models by Rasmusen et al. (1991) and Segal and Whinston (2000a), noting that the facts of the case were consistent with event promoters not internalising the contracting externality imposed on other promoters when signing up to exclusivity (an effect discussed in this chapter in Section 3.2.3.1).

Tying, Bundling and Bundled Rebates

4.1 Introduction

This chapter focuses on a relatively common business practice, that of selling two or more (different) products in combination. Products are often sold together in the marketplace. In fact, the distinction between a product and its components is often very blurred: cars come with a steering wheel (among many other features), wine bottles with a cork, smart phones with a battery. Firms can combine their sales in many different ways, as we describe next.¹

Pure bundling refers to the case in which a firm only offers the bundle as a package. Think of a hard drive, a keyboard, a screen and a touch pad all embedded in a laptop. Think also of a pay-TV contract offering a number of pre-packaged bundles of channels.

A firm, instead, engages in *tying* when it makes the sale of one of its products (the tying product) conditional upon the purchaser also buying some other products from it (the tied product(s)). A well-known example, that we will discuss in Section 4.7, has been Microsoft's former practice of selling its Windows operating system (the tying product) only in combination with the Internet Explorer (the tied product). Note that, differently from pure bundling, users could buy Internet Explorer as a stand-alone product, but could not obtain Windows without

¹ There are alternative definitions of the different forms that bundling can take. The nomenclature is somewhat secondary, once one is clear about the exact practice being considered.

Explorer. In this example, the two goods are sold in a fixed, one-to-one, proportion.²

In the case of variable proportions, sometimes known as *requirement tying*, it is left to the buyer to decide on the respective quantities. A frequent example is the requirement that purchasers of a firm's machine, say a printer or a copier, buy also consumables (for example, ink cartridges and toners) or after-sale services from the same firm.

Tying may be also equivalent to *full-line forcing*, that is, to the case in which a manufacturer supplies a product (or some products) to a buyer (say to a retailer) conditional on the retailer purchasing the whole range of products offered by that manufacturer.

Mixed bundling refers to the situation where a firm, besides offering the package of products at a given price, also supplies the individual products separately. Examples abound across industries. In the telecommunications and media industry, for example, many firms offer bundled packages of voice, internet access and TV (sometimes known as 'triple play'), but also sell these services independently. Some energy suppliers offer consumers bundles of gas and electricity (and sometimes car fuel too). A discount is usually available with the bundle, that is, the price of the bundle or package is typically lower than the sum of the prices of the individual products or services.

Sometimes such mixed bundling actually takes the form of multi-product rebates, or *bundled rebates*. These occur when a firm offers a discount conditional on the customer achieving sales (or sales growth) targets on a set of products. In a well-known US case, *LePage's*, which we will discuss more in detail in Section 4.6, 3M – the supplier of branded transparent tape with the Scotch product – offered rebates up to 2 per cent of total sales to customers (large stationery and grocery chains) meeting growth targets over six product categories (including stationery, home improvement and homecare products, for example). The higher the number of categories where the sales target was met, the higher the available rebate would be.

To conclude this classification of practices, it is important to distinguish between *contractual tying or bundling* and *technical tying or bundling*. The former simply involves the tie of two (or more) goods or services to customers or end-users; however, such goods remain distinct and the

² Despite those differences, in the rest of the chapter we will not make a distinction between pure bundling and tying, unless the context requires specifying the exact nature of the practice.

decision to combine their sales (and not to sell the individual components) can be easily reversed. Contractual tying (or bundling) can be replicated by mixed bundling, with the price for the individual components set at an arbitrarily high level such that demand for that component is zero. Also in this case, the decision to set such prices can be easily reversed. Technical tying (or bundling) instead involves the integration of the products (for example, in the design or in the manufacturing process) in such a way that it may be quite costly to undo the bundling or tying decision.³

Tying (or bundling) has a long history of scrutiny under US antitrust laws, and until the 1950s the US courts adopted a highly interventionist stance. The traditional argument underpinning such a severe treatment is known as the ‘leverage theory’: tying (or bundling) would provide a mechanism through which a firm with market power in one market can ‘leverage’ it to monopolise or gain market power in another market. Starting from the mid-1950s, the leverage theory has been heavily criticised by the group of scholars associated with the tradition of the University of Chicago, collectively known as the Chicago School. In this chapter, we will describe the traditional argument and the Chicago School’s challenge, followed by a review of more recent theories that have identified the specific conditions under which tying or bundling are likely to generate anti-competitive effects. As we did in the previous chapters, we will devote some attention to the models that share the main mechanism characterising this book, namely one based on scale economies.

This chapter proceeds as follows. In Section 4.2, we review the efficiencies that are often associated with these practices and we discuss the use of tying as a price discrimination device. In Section 4.3, we introduce in a relatively non-technical fashion the traditional argument on tying and bundling along with the Chicago School’s response, before turning to a discussion of some of the main mechanisms through which tying, bundling and bundled rebates may generate anti-competitive effects, based on a selection of economic models. In technical Section 4.4, we provide a formalisation of some of the key models discussed in the previous section. In Section 4.5, we seek to distil the key implications from the arguments and models reviewed in the previous sections and attempt to identify some workable criteria that may guide any antitrust intervention in this area. In Section 4.6, we discuss the landmark case-law from the US and the EU. In

³ As we will discuss in Section 4.3.3.1, this distinction may be relevant in the assessment of a potential theory of harm around possible anti-competitive effect flowing from tying: technical tying may involve a credible commitment to the tie for the future.

Section 4.7, we conclude the chapter with a more detailed discussion of a few antitrust cases investigated in different jurisdictions, linking them to some of the economic mechanisms presented in this chapter.

4.2 Efficiency and Price Discrimination Rationales

4.2.1 Efficiency Rationales for Tying and Bundling

Selling different products in a bundle is a common practice in the marketplace, across a number of industries with varying degrees of competition. One may expect this to be the case due to a number of potential efficiency gains that bundling brings about. Next, we discuss some of the categories of efficiencies that are typically associated with bundling.⁴

Scope economies in consumption Purchasing a bundle of products may give more value to a consumer than the purchase of the different components on a stand-alone basis. For instance, purchasing the bundle may reduce the costs of identifying the most appropriate combination of products that satisfy a complex need. And even if the most appropriate combination is known, purchasing a bundle may reduce the costs of actually looking for and purchasing each individual component (so-called ‘one-stop-shop’ benefits). Moreover, consumers may prefer to have the components already bundled rather than spending time assembling the components themselves.

Scope economies in production and distribution Producing two (or more) goods together may allow the manufacturer to exploit scope economies and reduce production costs. As long as it is impossible (or costlier) to sell the products separately once they are jointly produced, bundling is a natural strategy. For instance, Evans and Salinger (2007) document that the production of a tablet containing more than one active ingredient, say a pain reliever and a decongestant, involves lower costs than the production of two tablets, each containing an active ingredient. Combining different products may also generate a reduction in distribution and marketing costs (for example, subscriber acquisition costs; think for example of ‘triple play’ offers in the communications market, that is, the bundling of telephony, broadband access and pay-TV).

⁴ See also Evans and Salinger (2005) for a discussion of efficiencies arising in a few case studies.

Product innovation Another potential rationale for bundling is that a combination of products or services may have superior quality than the sum of the individual components. In a sense, some bundled products may be interpreted as innovations in and of themselves. For example, a modern car, complete with various state-of-the-art safety and communication technologies, is a better product than a car from the 1990s. In the *Microsoft* case (which we review in Section 4.7) Microsoft claimed that the technical bundling of the Windows operating system with the Internet Explorer browser, by intermingling their codes, gave rise to better performance than Windows with Internet Explorer installed at a later stage. Clearly, the extent of the product innovation benefits from bundling will depend on the specific circumstances of an industry, the technology used and the nature of the tying used by a firm.

Solution to reputation problems Bundling products together may also make consumers more certain about the quality of one of the components. Consider, for instance, a producer of copy machines. The user of the machine, if free to choose any provider of maintenance services, may come across one with inadequate expertise to repair that machine. As a consequence of poor maintenance, the machine may fail to work properly, with the user unable to identify the cause of malfunctioning (s/he may actually attribute it to poor quality of the machine). The equipment manufacturer may then suffer from an undeserved reputation loss. To avoid this problem, the manufacturer of the machine may decide to engage in requirement tying, by selling the equipment together with maintenance services (and possibly spare parts), thereby taking responsibility for the quality of repairs. In some cases bundling may be the only solution to the reputation risk, but in other cases alternative solutions may be found (for example, in most cases, suitable certification of qualified maintenance service providers is likely to constitute a valid solution).⁵

Efficient allocation under competition for slots Sellers of different portfolios of products often compete for limited slots belonging to a buyer. Think of the entertainment industry in which movies and series compete for slots on a television network; or of manufacturers competing for retail shelf-space. Jeon and Menicucci (2012) show that bundling, in this environment, by mitigating the intensity of competition that most valuable products suffer from less valuable ones, may favour an efficient allocation

⁵ See Schwartz and Werden (1996). See also Dana and Spier (2015).

of slots and may be welfare-beneficial. To see the intuition, consider the following example drawn from their paper.

There is only one buyer in a market – think of this buyer as a retailer, to fix ideas. The retailer has *two slots only* (that is, slots to only stock two types of products). There are two sellers: firm *A* with a portfolio of two products (product 1 and product 2), and firm *B* with a single product (product 3). The three products are independent (that is, they are not seen as complements or substitutes by end-consumers). Assume that product 1 and 2 generate a value of 40 and 30 (euros, for example), respectively, for the retailer, while product 3 generates a value of 20. Production costs are zero for all products.

The efficient allocation is such that the slots are occupied by the products of firm *A*, since these generate the greatest value to the retailer. However, under independent sales (that is, without bundling), the retailer has the possibility to fill up the slots with one product of firm *B* and one product of firm *A*. This exposes each product of firm *A* to competition for limited space from firm *B* and limits the revenues that accrue to firm *A*. Continuing with the numerical example, under independent sales firm *A* manages to sell both products only if it sets a price that does not exceed 20 for product 1 and a price that does not exceed 10 for product 2, earning a total profit that cannot exceed 30. To see why, consider that the retailer will buy both products 1 and 2 from firm *A* only if the surplus obtained from buying these two products ($40 - p_{A1} + 30 - p_{A2}$) is higher than both the surplus obtained by the combination of product 3 and product 1 ($20 - p_{A3} + 40 - p_{A1}$) and the surplus derived from buying the combination of product 3 and product 2 ($20 - p_{A3} + 30 - p_{A2}$). This is equivalent to saying that the surplus from buying product 1 and that from buying product 2 must be simultaneously higher than the surplus from buying product 3. That is, it must be that: $40 - p_{A1} \geq 20 - p_{A3}$ and $30 - p_{A2} \geq 20 - p_{A3}$. Since firm *B* will be willing to decrease its price p_{A3} down to zero in order to sell (given it has zero production costs), firm *A* cannot set a price p_{A1} higher than 20 and a price p_{A2} higher than 10, if it wishes to sell both products.

But firm *A* may have to decrease prices so much in order to sell both of its products that it may find it more profitable to sell just one. In the example, if firm *A* offers only product 1, it can set the price p_{A1} at 40 and it will make a profit equal to 40, which is higher than 30, the profit it would obtain from selling both products. This generates an inefficient outcome, since the retailer's preferred allocation (stocking the two products of firm *A*) does not materialise.

In this context, bundling removes the retailer's option to combine product 3 with one of the products of firm *A* and mitigates the intensity of competition felt by firm *A*'s products. This increases the revenues of firm *A* and restores its incentive to offer both products, thereby improving efficiency. Indeed, the retailer will buy the bundle as long as the surplus associated to it ($40 - p_{A1} + 30 - p_{A2}$) is higher than the one from buying only product 3, that is, ($20 - p_{A3}$); this – taking into account that firm *B* would be prepared to decrease its price p_{A3} down to zero – amounts to saying that the retailer will buy the bundle so long as $p_{A1} + p_{A2} \leq 50$. This gives firm *A* more profits and leaves the retailer with higher surplus than if firm *A* withdrew one product altogether.⁶

Note, however, that in the presence of budget constraints, the welfare effect of bundling might not be positive. As Jeon and Menicucci (2006) show, precisely by mitigating the intensity of competition suffered by a firm with a large portfolio of products, bundling may allow such a firm to exhaust the buyer's budget. Since the buyer has less money to spend on other independent products, this in turn reduces demand and the expected profitability of suppliers of such products. Bundling may then discourage entry or the activity of these suppliers.

Bundling in multi-sided markets A recent strand of the literature has analysed the potential efficiency-enhancing role of bundling in multi-sided markets, in which platforms provide goods or services to distinct groups of customers who seek to interact with each other. As also discussed in Section 1.2.3.2, examples of two- or multi-sided markets include payment cards, shopping malls, newspapers, yellow pages, pay-TV, web search engines, e-book readers, and so on. In many of these examples, there exist intergroup network externalities, whereby increased participation on one side increases the quality of the service enjoyed on the other side(s).

Multi-sided platforms need to have users onboard on all the sides so as to reach an efficient outcome and the price structure is used to solve such a coordination problem. In this context, bundling can increase efficiency by helping a platform implement the optimal price structure (that is, the optimal combination of prices across the various sides).

⁶ Note that also Whinston (1990) may be interpreted as a model of competition for slots. However, in that model efficiency dictates that one slot is occupied by a product of firm *A* and the other slot by the product of firm *B*. The inefficiency of independent sales identified above does not arise. Indeed, as shown in Section 4.3.3.1 it is bundling that may lead to an inefficient allocation by allowing firm *A* to win competition for both slots.

To see this point, consider a two-sided market with intergroup network externalities, in which a monopoly platform can find it optimal to set negative prices on one side in order to enhance participation and stimulate demand on the other side.⁷ Amelio and Jullien (2012) show that, if negative prices are not feasible, then bundling can be used as a tool to introduce implicit subsidies: the platform bundles either of the (two-sided) services and a good of particular interest for the targeted side, offering the bundle for free or at a discounted price. Think, for instance, of Google offering search services free of charge to web users; or of a magazine coming with a free DVD; or of free parking at shopping malls. Enhanced participation facilitated by bundling benefits consumers on both sides and increases the platform's profits.⁸

On a related point, Rochet and Tirole (2008) show that, when there is competition among platforms, bundling can increase welfare by helping the platform balance its price structure. The model is tailored to the specific case of the payment card industry, and allows users to multi-home, that is, to participate in different platforms. Indeed, consumers carry more than one payment card and merchants accept multiple cards.⁹ In the context of the payment card industry, the card association indirectly determines the cardholders' transaction fee and the merchants' fee by choosing the interchange fee, that is, the payment made by the merchant's bank to the cardholder's bank. Namely, the higher the interchange fee, the lower the cardholders' fee and the higher the merchants' fee. The optimal choice of such interchange fee must combine two objectives: it must be high enough so as to stimulate cardholders to use the card, but low enough to limit the merchants' resistance to accept the card. In this environment, Rochet and Tirole (2008) show that competition between card associations distorts the price structure. Since consumers hold multiple cards, then merchants have an incentive to reject the card that is most expensive for them. Competition results in the interchange fee being inefficiently low, which translates into the merchants paying too low a price and the cardholders too high a price, relative to the social optimum. Moreover, they show that, by tying a card

⁷ See, among the others, Armstrong (2006).

⁸ Amelio and Jullien (2012) show that the welfare effects are less neat if one considers different platforms competing with each other, because subsidising one side increases the opportunity costs of additional sales on the other side, thus generating a strategic effect and making competition softer on the profitable side.

⁹ Multi-homing is also common in other industries. For instance, users have more than one search engine or media player installed on their computer, and content providers offer content in more than one format.

facing competition (say a debit card) with another payment card that does not (or that faces less competition, say a credit card) an association gains flexibility to rebalance the fee structure. In fact, the two-card system can increase the interchange fee on the debit card, which translates into a lower debit cardholders' fee and promotes their use; at the same time the two-card system can lower the interchange fee on the credit card.

Choi (2010) focuses on the media software industry (where consumers and content providers are on either side of the platform) and identifies a different reason why tying can be welfare-enhancing. The critical ingredient of this model is that each platform (say each media software) has some exclusive content, that is, some content that is not suitable for the format of the other media software. In this context, multi-homing naturally arises. Imagine now that a platform monopolises another product or service that is necessary for consumers to participate in the two-sided markets – think of the operating system – and decides to tie the monopolised product and the platform service. The author shows that tying does not automatically foreclose the competing platform. Indeed, due to the existence of exclusive content, tying induces more consumers to multi-home. Moreover, tying allows consumers who would not have bought the tied platform independently to access the exclusive content that comes with it. As a result, total welfare increases. This model suggests that caution should be exercised when considering potential theories of harm based on market tipping and foreclosure when multi-homing (which is generally a form of increased competition) is a relevant market feature.

4.2.2 Tying and Bundling as Price Discrimination Devices

Some industries may present situations in which conventional forms of price discrimination are not possible. Tying and bundling have been identified as practices that allow a firm to restore its ability to profitably discriminate between consumers.

As we have already discussed in Chapter 2, the purpose of price discrimination is to differentiate the price paid by different consumers for a given product so as to accommodate differences in willingness to pay (or valuation) for a product or service across consumers. When explicit price discrimination is not feasible – for example because firms are legally prevented from engaging in price discrimination, or because firms know only the distribution of valuations across the population as a whole and

not the valuation of each individual – tying and bundling may represent an (imperfect) alternative.

Bundling (and in particular mixed bundling) is one alternative to explicit price discrimination. The scope for a firm to price discriminate via bundling is often driven by differences in consumers' valuations for a bundle being less pronounced than the difference in valuations for the individual products (the simple example below will clarify this statement; see also Section 4.3.3.2 for further discussion). This may allow the firm to extract part of the consumer surplus through the use of a single price.

This advantage of bundling comes at a cost, though: it may be that, with bundling, consumers having a very low valuation for one of the components, indeed lower than production costs, end up also purchasing that product, which sacrifices efficiency.

This drawback of bundling is mitigated by mixed bundling: by offering both the bundle and the individual components, the firm may combine the advantage of reducing heterogeneity in valuations with the possibility of targeting consumers with strong preferences for a particular component, thereby limiting efficiency losses and extracting more rents from consumers.

We will make use of a simple example to illustrate these points,¹⁰ before turning to the use of tying as a further alternative to explicit price discrimination and to some brief remarks on welfare implications.

Price discrimination through mixed bundling: An example Consider a monopolist offering two products, say mobile telephony services (product 1) and fixed line telephony services (product 2). The marginal cost of supplying each good independently is constant and equal to $c_1 = 18$ for product 1 and to $c_2 = 0$ for product 2. The marginal cost of supplying the two services in a bundle is just the sum of the component costs: $c_B = c_1 + c_2$. There exist three categories of consumers whose valuations for the individual products are indicated in Table 4.1.

¹⁰ This example is inspired by the seminal article by Adams and Yellen (1976). In a series of examples with an atomistic distribution of consumers, the authors show that a monopolist may increase its profits by using mixed bundling. This result was generalised by Schmalensee (1984) to a joint normal distribution and by McAfee, McMillan and Whinston (1989) to general distribution functions. See also Salinger (1995). Fang and Norman (2006) provide a generalisation to the case of n products, but they consider pure bundling only. Armstrong (1996) also provides a generalisation of optimal non-linear tariffs for a multiple-product monopolist, finding *inter alia* that some low-valuation consumers would not be served in equilibrium and that the optimal tariff is driven by the total cost of the bundle.

Table 4.1. *Consumers' valuations*

	Mobile	Fixed Line
Elderly People	5	95
Families	50	60
Young professionals	95	5
Unit Cost	18	0

Young professionals, who spend most of their time away from home, value significantly mobile telephony services while they do not care too much about having a fixed telephony at their place. Elderly people exhibit the opposite preferences. Middle-aged people with kids value both services moderately. Consumers' valuations for the bundle are given by the sum of the valuations for the individual components. For simplicity, we assume that there is one consumer for each category we described.

By assumption, the supplier cannot engage in open price discrimination. Then, if it offers the services independently – setting a uniform price for each service – the supplier's optimal choice will be to set a price $p_1^* = 95$ for mobile services, thereby selling to young professionals only, and to set a price $p_2^* = 60$ for fixed line services, thereby supplying elderly people and families. The total profits earned under uniform pricing amount to 197.¹¹ There is a deadweight loss in that there exist consumers having a valuation for a given service exceeding its production cost (that is, families for mobile services and young professionals for fixed line services) who are not served.

Consider next the case in which the supplier decides to bundle the two services. Preferences for the bundle are much more homogeneous than preferences for the individual components, with families having valuation 110 and both young professionals and elderly people having valuation 100. This increases the scope for rent extraction through a single price: the optimal bundle price, $p_B^* = 100$, allows the supplier to absorb the entire surplus of both elderly people and young professionals and to extract most of the surplus of the families. Profits increase to 246.¹² Note that now

¹¹ This is the sum of the profit from selling mobile telephony services to young professionals (with revenues of 95 minus costs of 18) and of the profit from selling fixed telephony services to each of the other two customer categories (with revenues of 60 per customer category and zero costs).

¹² That is, the firm earns revenues of 100 from each of the three customer categories and incurs costs for the provision of mobile services of 54, or 18 for each of the three customer

families have access to mobile services and young professionals to fixed line services, thereby increasing efficiency. The downside of bundling is, however, that elderly people have access to mobile services even though their valuation (equal to 5) falls short of marginal costs (equal to 18).

Mixed bundling (a menu of offers) can mitigate this flaw. The monopolist can offer both the bundle, at a price $p_B^* = 110$ (or, equivalently, mobile services in the bundle at a price $p_{1B}^* = 50$ and fixed line services in the bundle at a price $p_{2B}^* = 60$), and the individual components, at a price $p_1^* = p_2^* = 95$. Profits increase to 264¹³ because under mixed bundling the supplier can use a richer set of instruments to extract surplus from customers: it can use the price of individual components to target those who highly value only one product, while using the bundle's price to target consumers with moderate valuations for both products. Rent extraction is more complete ('less money is left on the table'), while avoiding forcing elderly people to buy mobile services (whose marginal cost exceeds these consumers' valuation). This is beneficial both to profits and efficiency. Some inefficiency still arises, since young professionals are excluded from access to fixed line services.¹⁴

Tying as a metering device Price discrimination, when consumers use a given product to different extents, can be achieved through other forms of

categories (recall that there are no costs in providing fixed telephony services in this example).

¹³ This is obtained by adding revenues of 110 from the bundle (bought by the family category), 95 from the sale of fixed telephony services to the elderly and 95 from that of mobile telephony services to the young professional category, and then subtracting the cost of providing mobile telephony services to the two customer categories purchasing them (36 in total).

¹⁴ The example helps illustrate another issue. The pricing structure that arises under mixed bundling is such that if one attributes to the mobile services the discount granted to the fixed line services in the bundle (that is, the difference between the individual price of fixed line services and their price in the bundle: $95 - 60 = 35$), the resulting effective price of mobile services in the bundle turns out to be $p_{1B}^{eff} = 50 - (95 - 60) = 15$, which lies below the marginal cost of mobile services ($c_1 = 18$). If the two products are not priced individually in the bundle, one obtains the same result by computing the difference between the price of the bundle and the stand-alone price of fixed line services (which is often denoted as the incremental price, in this case of mobile services): $p_B^* - p_2^* = 110 - 95 = 15 < 18 = c_1$. In this context where the monopolist faces no challenger (by assumption) and thus cannot have any exclusionary intent (since there is no one to exclude!), it is apparent that this pricing decision cannot be regarded as predatory or exclusionary. This example suggests that some care should be taken when comparing effective prices to costs so as to establish whether the use of mixed bundling is anti-competitive (see Sections 4.3.5 and 4.5 for a discussion of policy implications).

bundling, such as requirement tying (that is, tying is used as a ‘metering’ device).¹⁵ Think of the users of a printer: some, like students, may use it occasionally at home; others, like lawyers, may use it more intensively for their work. Different intensities of use translate into different willingness to pay for the printer (leaving aside the existence of printers of higher or lower specification). A manufacturer of printers would find it optimal to charge a different price depending on the intensity of use, say a low price to students and a high price to lawyers. If explicit price discrimination is not feasible (for instance because a lawyer delegates a student to buy the printer), requirement tying may allow the manufacturer to achieve an equivalent outcome.¹⁶ In fact, the manufacturer might set a uniform low price for the printer, but it might tie the equipment to consumables (for instance ink cartridges, toners, etc.) charging a relatively high price on each unit of consumable. In this way, lawyers, who use the printer intensively and need to buy more consumables than students, end up paying a higher price. (Note that for this strategy to be feasible, the producer must be able to be the only seller of consumables. We shall discuss this point further in Section 4.3.6 on after-markets.)

Welfare implications As for the welfare effects of bundling as a price discrimination device, the considerations already made in Chapter 2 continue to hold: the impact on welfare is generally ambiguous. Indeed, if by capturing more of a consumer’s surplus the firm is willing to serve a larger fraction of its potential market, then bundling may be welfare-beneficial by reducing deadweight losses. As discussed above, the drawback of bundling is that wider market participation may come at the cost of inefficiently supplying some components to consumers whose valuation is too low.¹⁷ However, this downside of bundling matters the less

¹⁵ This effect was already noted in Bowman (1957).

¹⁶ This is also an example of after-markets, which we discuss in more detail in Section 4.3.6.

¹⁷ Consistently with the ambiguous effects expected in theory, empirical studies on the welfare effects of bundling lead to mixed results. Luo (2013) studies the optimal selling strategy in a model that endogenises the bundling decision. Using data on internet and phone services sold by a major Chinese telecoms company, he estimates the structural parameters of the model. This allows the counterfactual scenario where bundling is not feasible to be identified. Unbundling is shown to decrease both profits (by 10.14 per cent) and consumer welfare (by 17.18 per cent). Crawford and Yurukoglu (2012) focus on the US cable TV industry. In their model the bundling decision is exogenous, but they incorporate the negotiation between downstream distributors and upstream channels. Keeping input prices fixed, they show that unbundling TV channels limits the scope for price discrimination, which as a result increases consumer surplus and decreases industry profits. However, in the unbundled scenario, input costs are higher, feeding into higher

the lower the marginal costs of production are. Thus information goods (such as news), whose marginal costs of production are close to zero, are particularly conducive to bundling and, in such a case, bundling is more likely to generate welfare gains.¹⁸

Importantly, price discrimination represents a possible rationale for bundling when a firm cannot engage in explicit price discrimination. Then, bundling is more likely to be used as a price discrimination device when a firm's customers are final consumers, as opposed to a context in which a firm sells to industrial customers or large retailers. In the latter case, prices and quantities are typically determined in bilateral negotiations, which creates the scope for explicit discrimination across buyers.

4.3 Anti-competitive Tying, Bundling and Bundled Rebates

In this section we will begin by referring to the case in which an incumbent firm is a monopolist in one market (which we will denote as market *A*) while it faces actual or potential competition in another market (denoted as market *B*). Within this setting, in Section 4.3.1, we first illustrate the criticism to the leverage theory (originally due to the Chicago School) in the case in which the anti-competitive concern raised by tying seems to be particularly severe, namely the extreme case where products are linked by perfect complementarity in consumption. We will then analyse the case in which the two products are independent. We will discuss next the different reasons why the criticism of the Chicago School may not apply. Economic theory has moved into three main directions to explain why it may be rational for a monopolist in one market to engage in anti-competitive tying or bundling.

One strand of the literature, which we discuss in Section 4.3.2, emphasises that, despite its monopoly position in market *A*, the incumbent may be unable to extract from consumers the whole surplus (or rent) available across markets *A* and *B*, either currently or in the future. For this reason (so-called imperfect rent extraction) tying may turn out to be more profitable for the incumbent than independent sales.

Another strand of the literature, which we review in Section 4.3.3, identifies the circumstances under which bundling allows the monopolist

prices and reducing consumers' gains from purchasing individual channels. Hence, in this scenario, unbundling is likely to be welfare-detrimental.

¹⁸ See Bakos and Brynjolfsson (1999) for a study on the strategy of bundling a large number of information goods.

in market *A* to commit to an aggressive behaviour which discourages entry in the adjacent, potentially competitive, market(s).

The third strand of the literature, which we discuss in Section 4.3.4, considers the possibility of entry in both markets *A* and *B*. One important paper in this literature, Choi and Stefanadis (2001), analyses the incentive to tie as a way to raise entry barriers. They consider a model with a potential entrant in each market and where each entrant, in order to be able to operate, needs its original investment in *R&D* to be successful, but this only happens with some probability. Another part of this literature argues that bundling can be an effective instrument to protect the core monopolistic position in market *A*. In particular, Carlton and Waldman (2002) show that the decision of the dominant firm to bundle monopoly product *A* with complementary product *B*, by reducing the *current* demand for the rival's product in market *B*, avoids *future* entry in market *A*. This is the case because the sales that the rival can make in the future, when it can also enter market *A*, are insufficient either to cover the fixed costs necessary to operate in both markets or to achieve the critical consumer base that makes its bundle of *A* and *B* superior to that of the incumbent. *Current* sales in market *B* are therefore crucial for the rival to achieve efficient scale and pave the way to future entry in market *A*. Bundling denies access to these critical sales and allows the incumbent to preserve its dominant position in market *A*. Then, the ability of the dominant firm to use bundling so as to exclude in a profitable way relies, as in the other chapters of the book, on the existence of important scale economies either on the supply-side or the demand-side.

When firms sell multiple products, they often have the ability to sell them in different combinations and at different prices. In Section 4.3.5, we therefore consider the potential for mixed bundling (and specifically, bundled rebates) to raise anti-competitive concerns under some conditions.

Finally, in Section 4.3.6, we set out some of the key aspects that are relevant to the assessment of tying in the context of after-markets, that is, markets where the purchase of a primary product (for example, a printer, a copier, or a car) is followed or may be followed by that of a secondary product (for example, a toner, maintenance or servicing) - a practice related to requirement tying.

4.3.1 The Chicago School Critique to the Leverage Theory

In what follows, we explain the Chicago School argument according to which a monopolist of a product may have the ability, but not the incentive,

to extend (or ‘leverage’) its dominance to another complementary product market by means of tying. Intuitively, by choosing appropriately the price of its products when it sells them independently, the incumbent will manage to extract more rents from consumers. For instance, it could price aggressively its competitive product so that consumers will have high surplus when buying the rival product, but appropriate that consumer surplus by charging a very high price for the monopolistic product.

To see this argument more rigorously, think of two products, A and B , and suppose that consumers enjoy utility (that is, value) U if they combine one unit of product A with one unit of product B , whereas consuming the products separately would not generate any value. Suppose also that a firm (‘the incumbent’) is a monopolist of product A , produced at constant marginal cost c_{IA} , while there exist several (actual or potential) competitors in the market for product B . For simplicity, assume that there are no entry costs to start operations in market B and that the incumbent’s rivals have marginal cost c_B , while the marginal cost of the incumbent is c_{IB} . The incumbent and the rivals produce homogeneous versions of product B and compete in prices. In this context, if the incumbent decides to bundle the two products (or to tie product A to product B), then an anti-competitive effect is automatically exerted. Consumers will have no incentive to buy product B from the incumbent’s rivals, unless these rivals give it away for free: they need product A to derive any value at all from B , and product A is sold uniquely by the incumbent together with B . So there will be no competition in market B and the incumbent will monopolise both markets by setting a price $\tilde{p} = U$ for the bundle. The margin that the incumbent would gain on each bundle is $U - c_{IA} - c_{IB}$.

The Chicago scholars do not challenge the fact that, once the incumbent bundles the two products, competition in market B is foreclosed. Their point is, rather, that the incumbent has *no incentive* to do so. Indeed, by selling the two products independently, the incumbent can make at least the same profits as in the case of bundling.

To see this, consider the case where products are sold independently. Since consumers’ willingness to pay for the combination of the two products is U , the highest price that the incumbent can set for product A is $p_{IA}^* = U - p_B^*$, where p_B^* is the equilibrium price paid by consumers in market B . This maximum price is the one that makes consumers indifferent between buying and not buying products A and B . Note that the lower the price of product B , the higher the price that the incumbent can set for product A . In this case in which the equilibrium price in market B

is $p_B^* = c_B$, the price for product A is $p_{IA}^* = U - c_B$. Then, suppose that the incumbent is less efficient than the rivals in the production of B: $c_B < c_{IB}$. Under stand-alone sales, the incumbent would not manage to sell product B. Yet, the margin that it makes on the sales of product A, $p_{IA}^* - c_{IA} = U - c_B - c_{IA}$, is larger than the margin obtained when it bundles the products and monopolises both markets, $U - c_{IB} - c_{IA}$. And the more efficient the rivals, that is, the lower c_B , the higher the price that the incumbent can set for product A and the higher its profit. If, instead, the incumbent is more efficient than the rivals in the production of B (that is, $c_{IB} < c_B$), then it would sell product B under stand-alone sales earning a margin $p_B^* - c_{IB} = c_B - c_{IB}$. Then the total margin of the incumbent under independent sales, $p_{IA}^* - c_{IA} + p_B^* - c_{IB} = U - c_{IA} - c_{IB}$, is the same as in the case of bundling.

This argument suffers from some limitations, which we will discuss in the next sections. Nevertheless, it raises an important point: when products display some complementarity (that is, when consumers value them more when bought in a combination), the anti-competitive effect of bundling is less obvious than it appears at first sight. Indeed, it is not necessarily the case that a monopolist of one component has an incentive to reduce the intensity of competition in the market for the other component. This is because it can benefit from the presence of rivals that have lower production costs or that produce differentiated versions of the second component. Thus, if a monopolist (or more generally a firm with market power) engages in bundling, its purpose is not necessarily to foreclose rivals from another market.

In the above discussion, the bundled products are complementary. The Chicago School's approach challenges the 'leverage theory' also when products are not complementary and generate utility when consumed on a stand-alone basis. Consider the same setting as above, with the difference that consumers are willing to buy at most one unit of good A, which they value v_A , and one unit of good B, which they value v_B . We show that, also in this case, it is never optimal for the incumbent to supply the bundle. Indeed, consumers are willing to buy the bundle from the incumbent if they obtain the same surplus as in the case in which they buy product B alone at the competitive price from the rivals. If the competitive price of product B is c_B , in order to persuade consumers to buy the bundle, the incumbent must leave at least the surplus $v_B - c_B$ to consumers. Then, the highest surplus that the incumbent can extract from consumers when selling the bundle is $v_A + v_B - (v_B - c_B) = v_A + c_B$. This surplus translates into a margin $v_A + c_B - c_{IA} - c_{IB}$ on each bundle. The incumbent would

make the same total margin by selling the two products separately – the sum of $v_A - c_{IA}$ from the sales of product *A* under monopolistic conditions and of $c_B - c_{IB}$ from the sale of product *B* in the competitive market - with the additional advantage that the incumbent could decide not to supply product *B* at all if its marginal cost c_{IB} exceeds that of the rivals, c_B . In other words, if the incumbent already extracts the *entire* consumer surplus from the sales of product *A* on a stand-alone basis, and it must leave consumers at least the same surplus as when they purchase product *B* on a stand-alone basis at a price equal to the rival's marginal cost, there is no additional surplus that it can extract by bundling the two products. Note that, for simplicity, we have referred to the simple case in which the demand functions of the two products are completely inelastic. If one assumes standard downward-sloping demand functions, the result that bundling is not profitable for the incumbent continues to hold as long as the incumbent manages to extract the entire consumer surplus when it sells the monopoly product *A* on a stand-alone basis, for instance because it can adopt two-part tariffs. (See the technical Section 4.4.1.1.)

4.3.2 Imperfect Rent Extraction

4.3.2.1 Imperfect rent extraction with complementary products

In the argument presented above, the incumbent has no incentive to engage in bundling (which would lead to the exclusion of more efficient competitors from market *B*). This is because it controls product *A* – which is essential for consumers to enjoy utility from the combination of product *A* and product *B* – and, through the price of the essential product, it can fully appropriate the additional surplus that more efficient producers of product *B* bring into the market. This argument, however, relies on the assumption that the incumbent is free to choose any price for product *A*. Imagine, instead, that the incumbent is subject to some constraint in the choice of p_A , for instance because that price has to be negotiated with some government or regulatory body (the ministry of health, as may be the case with some pharmaceuticals; a regulatory authority, if the price of product *A* is regulated). Then, under independent sales, the incumbent may fail to extract the whole surplus that consumers enjoy when they combine product *A* with the rivals' product *B* and it may find it more profitable to bundle the two products, thereby foreclosing more efficient competitors from market *B* and harming total welfare.

This argument is very similar to the one we present in Chapter 5, in which we analyse the incentives of a vertically integrated incumbent, which

is a monopolist in an upstream market and also competes in a downstream market, to exclude a more efficient rival from the downstream market. (Upstream production and downstream production can be considered as complementary activities.) In that case, the Chicago School argument is that the control of the essential input allows the incumbent, through an appropriate choice of the input price (or, more generally, of the contract offered to independent rivals), to extract the entire rents that a more efficient downstream firm is able to produce. For this reason, selling the input to the independent firm turns out to be more profitable than engaging in refusal to supply and serving the final market through the own, less efficient, subsidiary. However, as discussed in Chapter 5, the incumbent may fail to extract the whole rents from the independent firm, for instance because the sale of the essential input is subject to regulation (see Section 5.2.3.1). Then, the vertically integrated incumbent may find it more profitable to supply the downstream market through its own affiliate (which would be equivalent to the bundling decision discussed in this chapter) and to exclude the more efficient downstream rival.¹⁹

¹⁹ Let us highlight, though, a distinction between the setting analysed in this chapter and the one that we discuss in Chapter 5. In this Chapter the incumbent extracts the rents produced by more efficient suppliers of product *B* directly from final consumers, through the price it sets for product *A*, and not through contracts with its competitors in market *B*. Instead, in the setting analysed in Chapter 5, the vertically integrated incumbent extracts the rents produced by a more efficient competitor in market *B* through the conditions that it sets for the sale of the essential input to that competitor. For instance, the incumbent could optimally sell the input at marginal cost to the more efficient independent rival, let it monopolise the downstream market and then extract the monopoly profits through a fixed fee. This distinction matters because, as we also discuss in Chapter 5 (Section 5.2.3.2), in the vertical setting there is scope for opportunistic behaviour by the incumbent. This is another source of incomplete rent extraction that may lead to exclusion of the independent downstream rival. The intuition with respect to the vertical integration scenario is that, once the independent firm has committed to pay the fee that corresponds to monopoly profits in the final market, the vertically integrated incumbent has an incentive to earn additional profits by competing in the downstream market through the own affiliate. The anticipation of that behaviour on the side of the incumbent decreases the profits that the independent firm expects to make in the downstream market. Then, the independent firm is not willing to pay a fee that corresponds to the monopoly profits in the downstream market, unless the incumbent manages to credibly commit not to operate the own affiliate. However, this mechanism does not operate in the setting analysed in this chapter because the incumbent sells to consumers and not to firms in the complementary market and the scope for opportunistic behaviour does not arise in general (opportunistic behaviour may arise in after-markets, as we shall see in Section 4.3.6, but that will be in a very different context).

4.3.2.2 Imperfect rent extraction with independent products

Mixed bundling

We discussed above that bundling cannot increase the incumbent's profits when products are independent and the incumbent can extract the entire consumer surplus (or rent) from the sales of product A . Under downward-sloping demands, full surplus extraction requires the incumbent to engage in perfect price discrimination, for instance by adopting appropriate two-part tariffs. However, if full surplus extraction is not possible, for instance because the incumbent is constrained to use linear pricing, then bundling might end up being more profitable than independent sales.

The intuition is as follows. (We will provide the formal proof in Section 4.4.1.2. See also Greenlee et al., 2008.) Consider first the case in which the incumbent sells the two products independently. It will set the monopoly price p_A^m in market A . In market B the equilibrium price is c_B , which corresponds to the marginal cost of product B . The incumbent's total profits amount to the monopoly profits π_A^m from product A .²⁰

Consider now the case in which the incumbent decides to bundle the two products. Assume that the incumbent still sells product A on a stand-alone basis at the price p_A^m . In other words, the incumbent does not engage in pure bundling, but in *mixed bundling*, and it offers product A in the bundle at a price below the monopoly level and product B in the bundle at a price above its marginal cost c_B . Consumers can also obtain product B on a stand-alone basis at the competitive price c_B . By reducing the price of product A in the bundle from the monopoly level and by increasing the price of product B in the bundle from the competitive level, the incumbent can make a consumer indifferent (or just about better off) with the bundle relative to the case in which she purchases the two products on a stand-alone basis. This situation is depicted in the upper panels of Figure 4.1.

The market for product A is shown on the left, with quantity of product A on the horizontal axis and its price (and cost) on the vertical axis. The market for product B is correspondingly shown on the right. By rebalancing the prices of products A and B as just described, the incumbent increases the consumer surplus in relation to product A (since the consumer faces a lower price than previously) and decreases that derived from product B (since its price has gone up). Through a careful calibration of the

²⁰ For simplicity we are focusing on the case in which the incumbent is as efficient as rivals in the production of product B : $c_{IB} = c_B$.

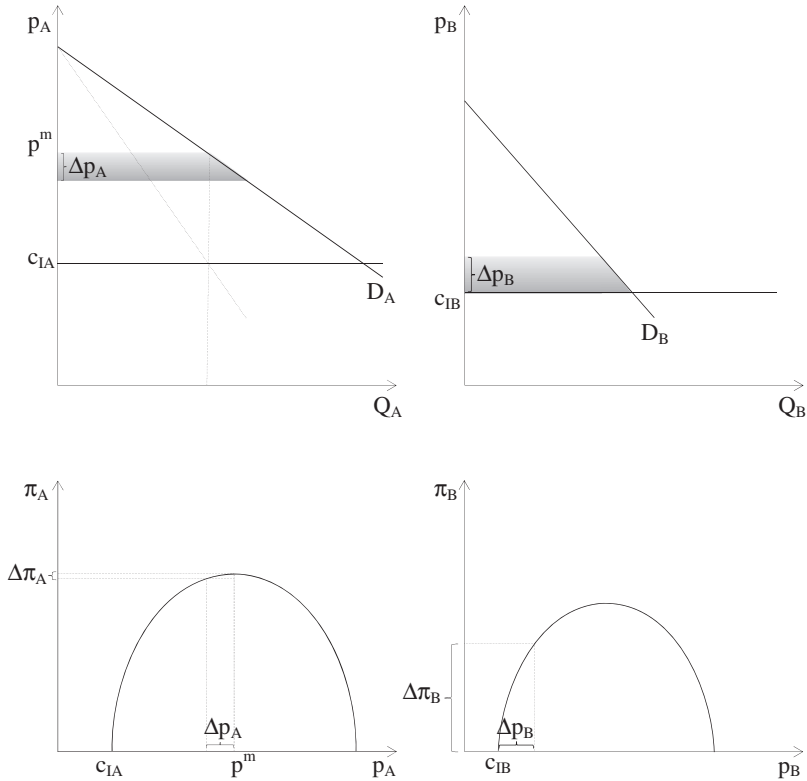


Figure 4.1. Mixed bundling as a profitable strategy, leaving consumers indifferent.

price changes, the incumbent can ensure that these opposite changes in consumer surplus offset each other (or the consumer is just about better off) – graphically, this is shown through the two shaded areas in the upper panels of Figure 4.1 being of equal size.²¹

Hence, the consumer will choose the bundle. However, the decrease in p_A from the monopoly level exerts a second-order reduction in the incumbent’s profits, while the increase in p_B from the competitive level exerts a first-order increase in profits. Hence, mixed bundling allows the incumbent to rebalance the prices of the two products in such a way as to increase its total profits relative to the case of independent sales.²² This

²¹ Mathematically, this rebalancing amounts to: $\Delta CS_A = CS_A(p^m - \Delta p_A) - CS_A(p^m) = CS_B(c_{IB}) - CS_B(c_{IB} + \Delta p_B) = \Delta CS_B$.

²² Note that bundling is welfare-beneficial in this case: consumers are as well off as in the case of independent sales, the incumbent makes larger profits and rivals in market B obtain zero profits in either case.

can be seen in the lower panels of Figure 4.1, which depict the incumbent's prices on the horizontal axis and its profits on the vertical axis (once again, product *A* is considered on the left and product *B* on the right). These lower panels show that the price rebalancing set out above leads to a relatively small decrease in the incumbent's profits from product *A* and a relatively large increase in profits from product *B*: graphically, the vertical distance $\Delta\pi_A$ on the left (that is, the loss in profits on product *A*) is shorter than the vertical distance $\Delta\pi_B$ on the right (the gain in profits on product *B*).

Pure bundling

In the case we have just analysed, the incumbent sells product *A* on a stand-alone basis *in addition* to selling it in the bundle. However, the incumbent could increase its profits further, relative to the case of independent sales, if it could *commit not to supply product A on a stand-alone basis*, that is, if it engaged in pure bundling (or full-line forcing, or if it committed to a prohibitively high price for product *A* on a stand-alone basis).²³ Essentially, by bundling the two products the incumbent uses the threat not to supply product *A* to induce the buyer to buy also product *B* at disadvantageous conditions.

To see the point, consider first the case in which the incumbent sells product *A* on a stand-alone basis. In that case, in order to induce the buyer to select the bundle, the incumbent must make her enjoy a consumer surplus that is slightly larger than the one enjoyed when she purchases the two products on a stand-alone basis and pays the monopoly price for product *A* and the competitive price for product *B*. In other words, the following constraint must be satisfied for the consumer to choose the bundle (where the superscript *b* denotes the price of a good sold in a bundle):

$$CS_A(p_A^b) + CS_B(p_B^b) \geq CS_A(p_A^m) + CS_B(c_B) \quad (4.1)$$

However, it is easier for the incumbent to persuade consumers to choose the bundle, if it commits not to supply product *A* on a stand-alone basis. In this way the alternative to the bundle becomes the consumption of product *B* only at the competitive price, and this alternative is clearly less appealing than consuming product *B* at the competitive price *and* product *A* at the monopoly price. The condition that has to be satisfied for the consumer to

²³ Greenlee et al. (2008) denote this latter case as 'bundled rebates', because product *A* stand-alone is sold at a prohibitive price while product *A* in the bundle is sold at a lower, therefore, discounted price. This practice is equivalent to pure bundling (refusing to sell a product alone or selling it at an arbitrarily high price have the same effect).

choose the bundle becomes:

$$CS_A(p_A^b) + CS_B(p_B^b) \geq CS_B(c_B). \quad (4.2)$$

As long as $CS_A(p_A^m) > 0$ – that is, as long as the incumbent is unable to extract the whole surplus from consumers through the monopoly price – the right-hand side of the constraint in equation (4.2) is smaller than the right-hand side of the constraint in the equation (4.1), and this gives more flexibility to the incumbent in the choice of the optimal prices.

Indeed, it is possible that constraint (4.2) is satisfied by choosing the monopoly prices p_A^m and p_B^m for the products in the bundle: this is the case as long as the gain in surplus that the buyer enjoys from consuming product A (at the monopoly price) in the bundle rather than not consuming it at all (a gain that amounts to $CS_A(p_A^m)$) dominates the surplus lost by paying the monopoly price for product B rather than the competitive price (a loss that amounts to $CS_B(c_B) - CS_B(p_B^m)$). When this condition is satisfied, selling the bundle is clearly more profitable than independent sales for the incumbent: bundling allows the incumbent to make the monopoly profits on both products rather than on product A only. Total welfare decreases because bundling causes a deadweight loss in the sale of product B .

When, instead, the gain in surplus from consuming product A (at the monopoly price) in the bundle rather than not consuming it at all is not that large, the incumbent must choose prices below the monopoly prices so as to satisfy constraint (4.2) and persuade the buyer to select the bundle. However those prices are still higher than the ones that the incumbent has to set so as to satisfy constraint (4.1). Then, also in this case, selling the bundle is more profitable than independent sales. The price of product A decreases in the bundle as compared to independent sales, whereas the price of product B increases. Consumer surplus overall decreases,²⁴ but this fall does not necessarily dominate the increase in the incumbent's profits. Then the effects of bundling on total welfare are unclear.²⁵

Credibility of the commitment not to sell product A on a stand-alone basis

The above discussion requires an important qualification. The possible

²⁴ When the threat not to supply product A on a stand-alone basis is credible, prices in the bundle are chosen so as to make the buyer enjoy surplus $CS_B(c_B)$ which is lower than the surplus $CS_A(p_A^m) + CS_B(c_B)$, which is what the buyer would obtain if she bought the two products independently.

²⁵ However, if one relaxes the assumption that market B is perfectly competitive, then bundling might exclude more efficient rivals from the market. In that case, the welfare losses caused by bundling would be larger.

welfare-detrimental effects of bundling (or full-line forcing) depend crucially on the ability of the incumbent to *credibly commit not to supply product A on a stand-alone basis* when it offers the bundle. Indeed, imagine the buyer does not select the bundle and decides to purchase only product *B*. In that case the incumbent has a strong incentive to make product *A* available (on a stand-alone basis) because it makes monopoly profits out of the sales of product *A*. Then, if bundling (or *full-line forcing*) consists only of a commercial decision, with the two products remaining distinct, the incumbent will reverse such a decision once the buyer decides to purchase only product *B* from the incumbent's rivals. A similar argument applies if bundling is achieved by setting a prohibitively high price for product *A* stand-alone, a decision that can be easily reversed. In other words, if it is not credible that the two products are offered only together, then the decision of the buyer is still governed by constraint (4.1) and the incumbent can use mixed bundling to rebalance the prices of the two products so as to earn higher profit, but without harming consumers (and total welfare) relative to the case of independent sales. Credibility can be achieved by technical bundling, which involves the integration of the products in such a way that it may be quite costly to undo the bundling decision; or it can be achieved by building a reputation of not selling product *A* on a stand-alone basis in an environment with repeated interactions. This discussion echoes the one we developed in Chapter 3 concerning exclusive dealing contracts and 'all-or-nothing' clauses. In that case the incumbent could profitably induce a buyer to purchase from it in exclusivity, thereby foreclosing a more efficient rival, threatening not to supply another 'must-have' product that the incumbent produces in case exclusivity is rejected. Also in that environment, such a threat facilitates acceptance of exclusivity if it is credible, a requisite which is not easy to be satisfied.

4.3.2.3 Imperfect rent extraction from future investments

Carlton and Waldman (2012) highlight another instance in which imperfect rent extraction may motivate the incumbent's decision to engage in bundling (or tying). (We provide a technical discussion of this model in Section 4.4.2.) Like in the standard setting of the Chicago School, they consider a firm (the incumbent) that is a safe monopolist in the primary market. The incumbent faces a more efficient rival in the secondary market. The primary product and the secondary product are complementary and they are durable, that is, they can be used for more than one period. To fix ideas, imagine that they can be used for two periods. The novelty of the model is that, in the second period, firms can invest so as to upgrade

the quality of the secondary product, or to develop a new complementary product. A key point is that investment generates additional utility (value) to consumers.

Absent the possibility to invest in the second period, the Chicago School argument would apply, and the incumbent would have no incentive to engage in bundling and exclude a more efficient rival from the secondary market. As discussed in Section 4.3.1, the lower the price of the secondary product, the higher the price that the incumbent can set for the primary product. Hence, the incumbent benefits from the presence of a more efficient producer of the secondary product because, through an appropriate choice of the price of the primary product, it can appropriate at least some of the additional surplus that the more efficient producer generates.

However, in Carlton and Waldman (2012) the incumbent, through the price of the primary product in period 1, cannot appropriate the additional surplus generated by the upgrading of the secondary product (or by the development of a new complementary product). This is the case because consumers are assumed not to be willing to pay in period 1 for future utility generated by an investment that has not taken place yet in period 1.²⁶

In this environment, the only way for the incumbent to appropriate the surplus generated by the investment is to discourage the more efficient rival from investing, so that investment is undertaken by the incumbent itself which will produce and sell the upgraded secondary product and will earn the associated profits in period 2.

Irreversible bundling allows the incumbent to discourage the rival's investment in period 2. This occurs because in Carlton and Waldman (2012), once the consumer has purchased the bundle from the incumbent, she cannot add the rival's secondary product to the system. In other words, the bundling decision does not only involve a joint sale of the two products, but also the impossibility for the rival to sell its secondary product in addition to the system (that is, primary and secondary product) sold by the incumbent. This may be rationalised by thinking that the incumbent does not only engage in tying but also denies the rival the information necessary to make its secondary product compatible with the incumbent's primary product, or that consumers would not be willing to have multiple versions of the complementary product (for instance due to high switching costs

²⁶ The dominant firm is also assumed not to be able to set a price for the primary product in period 1 that is contingent on whether upgrading will occur (that is, a contract with consumers providing, for example, that they will pay an overprice in period 2 if upgrading occurs, is ruled out).

driven by time, psychological or learning costs, or due to limited storage capacity). Whether this crucial assumption of the model holds in reality will depend on the context of the case at hand. Then, under bundling, the more efficient rival has no possibility to sell the upgraded complementary product in the second period, even though it is more efficient than the incumbent in its production, and it will decide not to invest.

It is important to note, however, that bundling also harms *the incumbent*. Under bundling, the more efficient rival cannot sell its secondary product in period 1 either, so the incumbent cannot appropriate (at least part of) the rival's efficiency rent through the price of the primary product.

Overall, bundling is profitable for the incumbent if the additional profits generated by upgrading its secondary product or by the development of a new complementary product are large enough compared to the efficiency rent of the rival which the incumbent forgoes when bundling. And precisely because it forecloses the activity of the more efficient producer of the secondary product, when bundling occurs, it is welfare-detrimental.

4.3.3 Bundling as a Commitment to Aggressive Pricing

In the argument of the Chicago School presented above, market *B* is assumed to be perfectly competitive. This makes it impossible to use tying or bundling in order to affect the structure of market *B*. Whinston (1990) departs from this assumption and introduces oligopolistic competition in market *B*. In this environment, his key insight is that bundling affects competition in the market. In particular, bundling may intensify price competition as compared to the case of independent sales, thereby limiting the profits that a firm expects to make in the market. In anticipation of this, bundling may discourage new entry.

In Whinston (1990) the reason why bundling makes price competition tougher is that it makes it more costly to lose a customer to the rival. The key idea is that, once the monopolist has committed to offering the two products only together, it can reap the monopoly profits from product *A* (over which it is a safe monopolist) only if it makes significant sales of the bundle. Differently stated, under bundling, losing a consumer to the rival reduces both the sales of product *B* and the sales of the monopolised product *A*. This induces the monopolist to adopt an aggressive competitive behaviour,²⁷ which limits the sales of the rival in market *B* and decreases

²⁷ In the terminology of Fudenberg and Tirole (1984), bundling commits to a 'top-dog' behaviour.

its profits below the level that would justify entry (or continued operation). To understand the logic of Whinston's argument, in Section 4.3.3.1 we first illustrate it with a simple example where the products are independent. Next, we will consider the case of complementary products, where Whinston's insight still applies, provided that product A is not essential in the eyes of consumers. We will then discuss the robustness of this result.

In a recent paper, Hurkens et al. (2016) highlight another reason why bundling may intensify price competition. The intuition is that bundling may reduce the heterogeneity in consumers' valuations as compared to the case of independent sales,²⁸ and for this reason it may increase the price elasticity of demand, thereby inducing firms to compete more aggressively. The novelty of their analysis is that, differently from Whinston (1990), bundling does not need to be irreversible so as to discourage entry. Moreover, as we will illustrate in Section 4.3.3.2, the level of dominance is a key element to assess the effect of bundling on competition.

Finally, Section 4.3.3.3 will go beyond pricing and will discuss cases in which bundling allows a monopolist to commit to an aggressive *R&D* behaviour, thereby discouraging rivals' innovation activity.

4.3.3.1 Bundling makes it more costly to lose a customer to the rival

Independent products To see Whinston's (1990) point, let us start from the case in which the two products are independent and let us go back to the above example where consumers value v_A one unit of good A and v_B one unit of good B . As before, the incumbent is a safe monopolist of product A , supplied at constant marginal cost c_{IA} , while there exists a potential entrant willing to start operations in market B . The entrant is more efficient than the incumbent in producing an identical version of product B ($c_{EB} < c_{IB}$), but it has to pay a fixed sunk cost f in order to start activity in market B .²⁹

Imagine that the entrant has already paid the entry cost and competes with the incumbent in market B . Let us consider first the case where the incumbent sells product A and product B separately. The incumbent sets the monopoly price v_A in market A , earning the monopoly margin $v_A - c_{IA}$. Instead, in market B , competition leads the incumbent to set a price at its marginal cost: $p_{IB} = c_{IB}$; the entrant slightly undercuts and captures the entire market. Note that, under independent sales, the incumbent has no

²⁸ For this same reason, bundling can be an effective tool to extract rents from customers when conventional forms of price discrimination are not possible. We discuss this issue in Section 4.2.2.

²⁹ For a more extensive formalisation, see Motta (2004: section 7.3.2.6).

incentive to set a price below its marginal cost in market *B*, as this choice would make it suffer losses in market *B*, without producing any beneficial effect on market *A*.

Assume now that the incumbent decides to bundle the two products. Imagine that the entrant chooses a price (a shade below) c_{IB} for product *B*, while the incumbent offers the bundle at a price $\tilde{p}_I = v_A + c_{IB}$. These offers lead to the same outcome that we set out under stand-alone sales, with the incumbent's bundle price being the sum of the incumbent's prices under stand-alone sales. There is an important difference though. Given these offers, consumers prefer to buy product *B* alone from the entrant rather than the bundle: in the former case they obtain a net value (utility) slightly larger than $v_B - c_{IB}$, while in the latter case their utility would be $v_A + v_B - v_A - c_{IB} = v_B - c_{IB}$. Since it fails to sell the bundle, the incumbent makes zero profits. Hence, it has an incentive to be more aggressive and decrease the bundle price further, so as to induce consumers to buy the bundle, thus making positive profits. It can be shown that at the equilibrium the entrant chooses $p_{EB}^* = c_{EB}$ and the incumbent sets the bundle price a shade below $v_A + c_{EB}$ and captures all the consumers. This price offer can be interpreted as if the incumbent implicitly set the monopoly price v_A for product *A* and a price slightly below c_{EB} for product *B*. Since the entrant is more efficient than the incumbent ($c_{EB} < c_{IB}$), the incumbent is implicitly selling product *B* below cost in the bundle. It has an incentive to do so because otherwise it would not sell the bundle at all and would lose the monopoly profits from product *A*.³⁰ The incentive to be so aggressive is, instead, absent when the incumbent sells the two products separately: it makes monopoly profits in market *A* irrespective of its ability to sell product *B*. Hence, it is willing to compete for market *B* as long as it does not suffer losses in this market.

The implication of the above analysis is that the decision to bundle the two products affects the rival's entry decision. When it observes that the incumbent has decided to bundle the two products, the entrant anticipates that the incumbent will be aggressive in pricing and that it (the entrant) will not be able to sell product *B*. So it anticipates that it would not be able to cover its entry cost. Instead, when it observes that the incumbent decided to sell the two products separately, the entrant anticipates that, if it enters the market, it will capture market *B* and will make positive profits.³¹ Therefore, by choosing to

³⁰ Of course, the underlying assumption is that monopoly profits from market *A* are larger than the loss suffered in market *B*: $v_A - c_{IA} > c_{IB} - c_{EB}$.

³¹ This example is extreme because when the incumbent bundles the two products, the entrant does not sell at all. More generally, by making the incumbent more aggressive

bundle the two products, the monopolist in market *A* deters entry in market *B* and ‘leverages’ its monopoly power from one market to the other.

A crucial aspect to note is that the incumbent’s ability to profitably use bundling so as to ‘leverage’ its market power depends on whether it can *credibly commit* to its commercial strategy in advance, that is, on whether it is impossible (or too costly) to undo the bundle if entry (for whatever reason) actually occurred. Suppose that bundling involves a marketing decision that can be easily changed (contractual bundling). In this case, the entrant will enter the market even if it observes that the incumbent decided to bundle the products. This is because it anticipates that, once entry occurs, the incumbent will decide to undo the bundle and sell the two products independently, because it would be in *I*’s interest, at that point in time, to do so. This occurs because bundling, by inducing the incumbent to be aggressive in pricing, reduces the incumbent’s profits (as well as the entrant’s), so the incumbent would be irrational in sticking to its bundling strategy if entry occurred; instead, it would prefer softening competition (by unbundling).³² Suppose instead that, in a different context (or industry), bundling is accomplished through product design (technical bundling). It may therefore be quite costly to undo the bundling decision once it has been taken, quite possibly costlier than enduring low or small negative margins if continuing to bundle in the presence of a rival. Anticipating this, the potential entrant will optimally decide not to enter, making the bundling strategy profitable for the monopolist, who would be successful in foreclosing potential competition. In sum, the closer the bundling decision to being completely irreversible, the stronger its value as a commitment device for the monopolist and the likelier that exclusion will take place, all else equal.³³ Note also that the bundling strategy is less likely to be chosen the more differentiated the products of the rival and the more efficient it is, since in these circumstances the rival will be more profitable and hence – *ceteris paribus* – less likely to be excluded.

in pricing, bundling limits the sales of the entrant/rival. If this effect is sufficiently strong, entry will not be profitable.

³² In the example made above, when entry takes place, the incumbent’s profit is $v_A - c_{IA}$ without bundling and $v_A + c_{EB} - c_{IA} - c_{IB}$ with bundling. The latter is lower because $c_{EB} < c_{IB}$ and the incumbent makes losses on product *B*.

³³ Of course technical bundling may also generate significant efficiencies, as discussed in Section 4.2.1. This makes it somewhat difficult to draw clear policy implications with respect to the use of technical tying.

Complementary products where product A is not essential When products are complementary, the Chicago School argument relies on the fact that the lower the price paid by consumers for product *B*, the higher the price that the incumbent can set for product *A*. For this reason, the incumbent benefits from the development of competition in market *B*. A crucial assumption for the existence of this linkage between the two prices is that good *A* is produced uniquely by the incumbent and is essential for consumers to derive value from the joint consumption of *A* and *B*. Imagine, instead, that there exists an inferior alternative to product *A*. The existence of this product constrains the price the incumbent can charge for component *A*. The incumbent might therefore prefer that no competition develops in market *B*.

To see this point, suppose that the inferior alternative, call it product A' , generates lower utility (gives consumers less value) than *A* when combined with product *B*, namely utility $U - \gamma$. Product A' is competitively supplied and its unit cost is the same as the incumbent's. Then, under stand-alone sales, the maximum price the incumbent can set on product *A* is the one that makes consumers indifferent between the combination $A + B$ and the inferior combination $A' + B$:

$$U - p_{IA}^* - p_B^* = U - \gamma - p_{A'}^* - p_B^* \quad (4.3)$$

where $p_{A'}^*$ is the (equilibrium) price of the inferior component A' . This leads to $p_{IA}^* = \gamma + p_{A'}^*$. In this case where $p_{A'}^* = c_{IA}$, the highest price that the incumbent can set on *A* is $p_{IA}^* = \gamma + c_{IA}$ and its profit margin on product *A* is γ (which can be interpreted as a premium for selling a superior version of the product).

Note that the competitive pressure exerted by the inferior component A' removes the linkage between the price of component *B* and the price that the incumbent can set on product *A*, thereby preventing the incumbent from increasing the price of component *A* as product *B* becomes cheaper. Thus, the incumbent does not necessarily benefit from more competition in market *B*. Indeed, the presence of the inferior product uncouples the two markets, making this scenario more akin to the one where products are independent and the profit margin that the incumbent can make in market *A* is fixed (and equal to $v_A - c_{IA}$). Similarly to what we have discussed for the case of independent products, bundling is irrelevant for the incumbent if market *B* is perfectly competitive. If, instead, one allows for oligopolistic competition in market *B*, then the incumbent may have an incentive to resort to irreversible bundling so as to credibly commit to an aggressive

pricing policy and deter entry in market *B*. Lack of competition in market *B* would allow the incumbent to act as monopolist on both markets and would make bundling a profitable strategy.

Robustness In some circumstances bundling, rather than intensifying competition in the market for product *B*, may make it softer. This effect may arise if one assumes that different consumers have a different willingness to pay for product *A*.³⁴ To see the intuition, consider first the case in which the two products are sold separately. The incumbent and the rival compete head-to-head in market *B*: they sell the same product and each firm can win the entire market from the rival by slightly undercutting it. Instead, when the incumbent sells the bundle and the rival sells product *B* on a stand-alone basis, it is as if vertical differentiation is introduced in market *B* (that is, different qualities are introduced). Indeed, all consumers value the bundle more than product *B* on a stand-alone basis, but they differ in how much more they value the bundle. In this environment, by decreasing its price for product *B*, the rival does not manage to capture the entire market: it will attract those consumers that do not care too much about product *A* and thus prefer to only buy product *B*, whereas those consumers having a high willingness to pay for product *A* would prefer to buy the bundle from the incumbent. In other words, the rival has weaker incentives to undercut as compared to the case of stand-alone sales. This, in turn, induces a softer response by the incumbent, which may counteract the ‘Whinston’ effect discussed above, leading to higher prices and higher profits relative to the case of stand-alone sales. As a consequence, bundling *attracts* entry rather than deterring it.

The fact that bundling makes it difficult for a single-product rival to win market share from the incumbent is the reason why, in Nalebuff (2004), it can facilitate entry deterrence. However, this result crucially hinges on the assumption that the price chosen by the incumbent – before entry takes place – has commitment value, that is, it cannot be modified. In such a context, the price that the incumbent must choose so as to prevent entry is higher under bundling than with stand-alone sales, precisely because for given prices the rival sells less under bundling than with stand-alone sales. Consequently, the incumbent’s profits are larger in the former case. However if, as it may be reasonable to assume, the incumbent can adjust its price in response to entry, then the entry

³⁴ This possibility is accounted for by Whinston (1990). An extensive analysis can be found in Carbajo et al. (1990), Chen (1997) and Nalebuff (2004).

deterrence effect breaks down. As explained above, bundling would make post-entry price competition softer, thereby making entry more profitable.

4.3.3.2 Bundling increases the price elasticity of demand

Hurkens et al. (2016) highlight another reason why bundling affects the way firms compete in the market and, as a consequence, the scope for entry. In their model there are two firms, each producing a different variety of both product *A* and product *B*. Consumers have heterogeneous valuations for each product. Valuations for product *A* and for product *B* are independently distributed. To fix ideas, assume that each distribution function is uniform over a given interval and that the two firms are located at the extremes of such interval.³⁵ Additionally, in each market, the model allows one variety to be superior in terms of quality to the other.

Let us start from the case in which the two products are sold independently and let us assume that the incumbent is the high-quality producer in market *A*, whereas the rival is the high-quality producer in market *B*. Hence, in market *A*, equilibrium prices are such that the market share of the incumbent is above 50 per cent, that is, it dominates the market.³⁶ This means that the dominant firm captures also the consumers that are 'located' close to the mean (that is, close to the middle of the interval), and leaves to the rival the consumers who value poorly its product and are 'located' close to the opposite extreme.

Next, consider the case in which each firm sells a bundle. Imagine that the incumbent's quality advantage in market *A* is large enough to make the incumbent's bundle superior to the bundle of the rival. The key insight of the model is that bundling makes consumers *more homogeneous*: the distribution of valuations for the bundle is more concentrated around the mean as compared to the distribution of valuations for individual products, meaning that there are more consumers whose valuation for the bundle is close to the mean and fewer consumers with extreme valuations.³⁷ The 'homogenisation' effect has two implications. First, for given prices – set

³⁵ The model is more general and considers any log-concave distribution which is symmetric around the mean.

³⁶ The model encompasses the special case in which the quality gap in market *A* is so large that the incumbent is a safe monopolist.

³⁷ Technically speaking, the distribution of the average valuation for the bundle is more peaked around the mean than the distributions of the valuations for the individual products. To see this, imagine that valuations for product *i*, with $i = A, B$, have three possible realisations, 0, 1/2 and 1, each one with probability 1/3. The average valuation for the bundle, $(v_A + v_B)/2$, has instead five possible realisations: 0 with probability 1/9,

at the sum of the (equilibrium) prices of the individual components – consumers’ homogenisation increases the demand of the dominant firm and decreases the demand of the rival: since the incumbent dominates the bundle market, it attracts also the consumers whose valuations are close to the mean, and these consumers are more numerous than under individual sales. Second, the ‘homogenisation’ effect changes the elasticity of demand and therefore the equilibrium prices. However, the way demand elasticity changes depends very much on the extent of the incumbent’s dominance.

If dominance is limited, at prices equal to the sum of the individual prices, the incumbent’s market share is slightly above 50 per cent. This means that firms compete for consumers located close to the mean who, under bundling, are more numerous as compared to the case of individual sales. As a consequence, a price cut attracts more demand and is more profitable relative to that case: bundling intensifies competition. Then, both the ‘demand size’ effect and the ‘elasticity’ effect are detrimental to the rival, while they operate in opposite directions as far as the dominant firm is concerned. However, when dominance is limited, the ‘elasticity’ effect is stronger than the ‘demand size’ effect, and bundling decreases not only the profits of the rival but also those of the dominant firm. In this case the same conclusions as in Whinston’s model apply: bundling, by making expected competition tougher, discourages entry of a new rival. However, since it decreases also the profits of the incumbent firm, bundling needs to be irreversible in order to act as an entry barrier.

If, instead, dominance is very pronounced, at prices equal to the sum of the individual prices, the incumbent’s market share is far above 50 per cent. This means that firms compete for consumers located far from the mean who, under bundling, are scarcer than in the case of individual sales. Hence, a price cut attracts fewer consumers and is less profitable relative to that case: bundling softens price competition. Both effects are beneficial to the dominant firm, while they operate in opposite directions as far as the rival is concerned. However, when dominance is very large, the ‘elasticity effect’ dominates the ‘demand size effect’ and bundling increases not only the profits of the dominant firm but also those of the rival. In this case bundling may be detrimental to competition not because it discourages new entry (in fact, it encourages new entry) but because it acts as a ‘competition-softening’ device among active firms.

Finally, for intermediate levels of dominance, bundling decreases the profits of the rival while it increases those of the dominant firm. Bundling discourages new entry also in this case, but it is important to notice that it does not need to be irreversible so as to exert an anti-competitive effect, because the dominant firm has an incentive to stick to such a practice even if entry should occur.³⁸

4.3.3.3 Bundling as commitment to aggressive *R&D*

We have just seen how the decision to bundle two or more products can act as a credible commitment for an incumbent to price aggressively conditional upon the entry of a rival. However, the same mechanism can take place in relation to strategies other than price-setting. In particular, Choi (1996, 2004) shows that the decision to bundle two products represents a credible commitment to an aggressive *R&D* behaviour, which may allow the incumbent to extend its dominant position to other markets.

Choi rationalises this effect in two ways. In Choi (2004), the incumbent and the rival invest in cost-reducing *R&D* ('process innovation') in market *B*, before they compete on price. As discussed above, Whinston (1990) shows that, under some circumstances, (irreversible) bundling induces the incumbent to be more aggressive in its pricing as compared to the case of stand-alone sales. This allows the incumbent to capture a larger share of market *B*, while limiting the sales of the rival. Then, when innovation decisions are taken, the incumbent anticipates that, by bundling, it will benefit more from a given amount of cost reduction than in the case of stand-alone sales, because such a cost reduction will be spread over a larger number of units sold. This will induce the incumbent to invest more in *R&D*, leaving the entrant to anticipate that it would sell less upon entry, thus reducing its incentive to invest or innovate. The rival's activity in market *B* will therefore be marginalised. Choi (2004) finds that tying reduces total welfare, but acknowledges that this result impinges on a series of assumptions made in the model.³⁹

³⁸ Peitz (2008) builds a model, based on Whinston (1990), with product differentiation and consumer preferences distributed two-dimensionally. Using a specific numerical example, Peitz (2008) shows that one can find examples where (i) bundling is profitable even once entry has taken place (so there is no concern about the credibility of a commitment to bundle); (ii) bundling discourages entry (for sufficiently high fixed costs of entry); and (iii) welfare is lower under monopoly (that is, when entry is deterred) than under independent pricing (that is, in the absence of bundling).

³⁹ In particular, whether tying leads to a reduction in total welfare in Choi (2004) depends on: the extent of fixed costs (or, put otherwise, scale economies) of *R&D*; whether there

In Choi (1996) the incumbent and the rival invest in *R&D* both in market *A* and in market *B*, with the two products being complementary. In each market, the outcome of the *R&D* race (or competition between firms) depends on the initial (pre-*R&D*) asymmetry between the two firms. The more efficient firm (that is, the one with lower *ex ante* production cost or better product quality) benefits more from a given amount of cost-reducing (or quality-enhancing) investment. Then, it is willing to bid more (invest more) for the innovation, thereby winning the innovation race and remaining ahead of the rival. Suppose next that the incumbent has a *large* initial advantage in market *A*, while the rival has an initial advantage in market *B* but such an advantage is *small*. When products are sold separately, the *R&D* races develop independently in the two markets. The incumbent wins the innovation in market *A* while the rival is successful, and becomes a monopolist, in market *B*. By contrast, when the two products are bundled, the innovation race at the product level is replaced by a single innovation race across the two products together. In this case, if the incumbent's advantage in market *A* is strong, while the rival's in market *B* is weak (that is, there is a form of incumbency advantage), the incumbent is (by assumption) more efficient than the rival (before innovation takes place) over the combination of products *A* and *B*. This provides it with an incentive to invest more, outbidding the competitor in the *R&D* race over *A* and *B* together, thus becoming a monopolist in both markets. In other words, the pooling of the two markets under bundling allows the incumbent to shift the slack in one *R&D* market to the other, thereby excluding the rival.

4.3.4 Bundling where Entry is Possible in All Markets

When products are complementary, the Chicago School critique relies on the assumption that market *A* is a 'safe' monopoly for the incumbent. By relaxing this assumption, one can rationalise the use of bundling (tying) so as to exclude rivals. We have already discussed the argument proposed by Whinston (1990) that assumes the existence of an inferior alternative to the dominant firm's product *A*. In this section, we review two further mechanisms through which bundling can generate anti-competitive effects,

is uncertainty in the outcome of the *R&D* investment; and on the extent of product differentiation between the incumbent and the entrant.

one due to Choi and Stefanadis (2001) and the other due to Carlton and Waldman (2002).

4.3.4.1 Bundling where entry is possible in all markets *simultaneously*

Choi and Stefanadis (2001) assume that the incumbent faces the threat of entry across all of the bundle's components, that is, there exists a potential entrant both in market *A* and in market *B*.⁴⁰ The model mainly refers to dynamic industries where the primary mode of entry is via innovation, with substantial upfront *R&D* spending. In each market entry occurs if the entrant's innovation is successful and produces a superior technology. Such technology allows the entrant to produce the same good as the incumbent at lower cost (or equivalently to produce a superior variant of the product). The larger the *R&D* investment, the higher the probability of success. In this context, absent bundling, the incentives to invest in either market are independent of what happens in the other market. Each entrant anticipates that it will be able to sell its product if its innovation process is successful, and will decide its optimal investment level accordingly by comparing the marginal cost and the marginal benefit of investment. When, instead, the incumbent decides to bundle the two products, it is not sufficient for an entrant to be successful in its own innovation process for it to gain access to consumers and sell its own product. This entrant also requires the potential entrant in the other market to be successful in its own innovation process so as to make the complementary product available to consumers. As a result, bundling makes the benefits of investment more uncertain, thereby weakening investment incentives and potentially generating consumer harm (both through softer price competition due to the exclusion of competitors and to lower innovation levels). It is also easy to see how the larger the number of products in a bundle (that is, of markets), the more uncertain the benefits of innovation for a given entrant, given it would need to rely on all other entrants being successful in their own markets.⁴¹

Note that the decision to bundle the products exerts two opposite effects on the incumbent's profits. On the one hand, by reducing the entrant's investment, bundling reduces the probability that both innovation processes are successful, thereby making it less likely that entry occurs in both

⁴⁰ The main part of the model assumes that the entrants are two distinct firms. The paper shows that the results extend to the case where the same firm tries to enter the two markets.

⁴¹ See Motta (2004: section 7.3.2.7) for a formal illustration of this model.

markets and that the incumbent is displaced by the more efficient entrants. On the other hand, bundling makes it impossible that entry occurs only in one component market, in which case, as highlighted by the Chicago scholars, the incumbent benefits from the presence of a more efficient rival by increasing the price of the complementary component under its control and appropriating (part of) the efficiency gains that the new producer brings into the market. The incumbent will opt for bundling when the risk of being displaced in both markets dominates the benefits of entry in a single market.

Note also that two crucial factors underlying the argument proposed by Choi and Stefanadis (2001) are that (i) the bundling decision is taken by the incumbent before the entrants make their investment decision and that (ii) such a bundling decision is irreversible. In this model bundling stifles the entrants' investment incentives because in the scenario in which one entrant is successful in its innovation process but the other is not, the successful firm cannot benefit from its investment because it cannot reach consumers and sell its product. Anticipating this, the entrant will decide to invest less. However, this scenario in which entry occurs only in one component is the one that benefits the incumbent through the partial appropriation of the rival's efficiency gains. Then, if the bundling decision could be changed, the incumbent would unbundle and the *ex ante* effect on the entrants' investment decisions would disappear.

4.3.4.2 Bundling to defend the monopoly market

In the model just reviewed, entry is possible in both markets simultaneously. Instead, Carlton and Waldman (2002) assume that initially only market *B* (denoted as the secondary market) is under the threat of entry, while the incumbent's position in market *A*, denoted as the primary market, is safe. At some future point in time, entry will be possible also in market *A*. The two products are complementary and, crucially, they are characterised by scale economies. (This model will be analysed more formally in Section 4.4.)

In a first variant of the model, scale economies arise on the supply-side, that is, there exist entry costs that the rival has to pay in order to start its operations and that the incumbent has already sunk. In this context, bundling the two products makes it impossible for the entrant to sell product *B* in period 1 – that is, in the period in which the incumbent's position in the primary market remains unchallenged – and to collect revenues from the sales of this product. The entrant could collect revenues

in period 2, when it can also enter the primary market and reach consumers despite the incumbent's bundling decision. However – and this is a crucial assumption of the model – the revenues earned in the second period alone are insufficient to cover the fixed costs necessary to enter the two markets, whereas such costs are covered if the entrant earns revenues in market *B* in the first period as well. Therefore bundling, by denying the entrant access to such crucial rents, deters entry altogether, that is, first in the secondary market (market *B*) and then in the primary market (market *A*).

A similar mechanism arises in the second variant of the model, in which scale economies arise on the demand-side and are due to the presence of network externalities in the consumption of the secondary product. Put otherwise, if the entrant manages to sell its secondary product both to the first and to the second cohort of consumers, then its secondary product joined with the primary one would exhibit superior quality to the incumbent's system (defined as the combination of its primary and secondary products). Instead, if the entrant does not sell its secondary product in the first period, then the sales to the second cohort of customers alone are insufficient to achieve superior quality. As a consequence, in this latter case, entry in the primary market is pointless for the entrant. Thus, also in this variant of the model, bundling deters entry altogether by denying the entrant access to (crucial) sales in the first period.

With regard to this (network effects) variant of Carlton and Waldman's model, we note that the same effects as bundling can be achieved by engaging in predatory prices in the complementary market: the purpose of the incumbent is to prevent the rival from selling the complementary product in period 1, so as to prevent the rival from achieving the critical customer base it needs to operate successful. The incumbent can adopt either bundling or low prices in market *B* so as to prevent the rival from selling. When, instead, the purpose of the incumbent is to deter entry, bundling cannot be replaced by predatory pricing.

A first important remark is that in Carlton and Waldman (2002) the reason why the incumbent has an incentive to choose bundling is not to deter entry in, or 'leverage' market power to, the secondary market *by itself*. Indeed, the incumbent would benefit from the entry of a more efficient supplier in the secondary market (as it would extract some of the surplus the latter brings into the market), as long as its monopoly position in the primary market is unchallenged. The key point of this model is that entry in the secondary market in the first period paves the way to future entry in the primary market. This generates the incumbent's incentive to choose

bundling as a defensive strategy that enables the incumbent to protect its dominant position in the primary market.

Thus for this model to be applicable (that is, for a bundling decision to be potentially challenged as anti-competitive consistently with the model), it is essential to show that there exists a linkage between immediate entry in the secondary market and the profitability of future entry in the primary market. In the model such a linkage comes through the existence of either fixed entry costs or of demand-side externalities in the secondary market. Whatever the origin, the important insight of this model is that there must be a reason why without immediate entry (or expansion) in market *B* the entrant would not find it profitable to enter (or expand) market *A* in the future.

A second remark is that, by bundling, the incumbent sacrifices profits in the first period. Bundling is a profitable strategy for the incumbent if the benefit from preserving its monopoly profits in the future dominates the first period cost. This consideration suggests why, in the variant of the model where scale economies arise on the supply-side, for the anti-competitive effect to arise, the bundling decision must be taken before the rival takes its entry decision in the secondary market and such a bundling decision must be *irreversible*. In fact, if bundling could be easily removed, the entrant would immediately enter market *B*, irrespective of the incumbent's bundling decision. It would anticipate that once entry in the secondary market has occurred, the incumbent would prefer to undo the bundling so as to appropriate some of the entrant's efficiency advantage. The entrant would then earn revenues in market *B* in the first period and would find it profitable to also enter the primary market in the second period.⁴²

Finally, note that the reason why bundling is an effective instrument to exclude a rival in this model is the one common to the different chapters of this book. Due to the existence of scale economies (either on the supply- or the demand-side) collecting rents (or selling) in the secondary market in the first period is key for the entrant to achieve efficient scale and to make its overall entry strategy (in the secondary market and in the primary market) profitable. The instrument used by the incumbent to deny the rival access to such crucial rents (or sales) is, according to this model, the bundling of product *B* and the complementary product *A*.

⁴² The formal analysis in Section 4.4 will clarify why irreversibility of the bundling decision is not crucial in the network externality variant of the model.

4.3.5 Anti-competitive Bundled Rebates

According to the economic literature that we have discussed so far the practice that may generate anti-competitive effects is the dominant firm's decision to sell two (or more) products together (bundling), or the decision to condition the supply of a product to the purchase of another product (tying). The prices associated to the bundle are not the source of the exclusionary effect.^{43,44}

In this section, we focus instead on the case in which the dominant firm offers a discount conditional on the buyer purchasing a bundle of products (or, more generally, achieving a pre-determined target concerning the sales of a set of products). Consider for instance the case in which the incumbent is a safe monopolist in market *A* and it offers product *A* at the monopoly price; moreover it offers a discount if the buyer also purchases product *B* at a price p_B . The discount may consist of a lump-sum discount or of a per-unit discount on the price of product *A*. As a consequence of this price offer, a buyer that purchases product *B* from a rival of the dominant firm loses the discount. Therefore it is not enough for the rival to undercut p_B so as to attract the buyer's purchases: it needs to offer a price that is sufficiently lower than p_B to compensate the buyer for the lost discount. In this case, bundled rebates essentially represent a way to implement a low price of product *B*.

Indeed, the price of product *B* implemented by the dominant firm may be so low that a more efficient rival is excluded from the market. If so, the reason behind this bundled rebate may be found among the theories of predation discussed in Chapter 1. It may be the case that market *B* features scale economies that the dominant firm has already exploited, while the rival has not (like in the model introduced in Section 1.2.3 in the context of predatory pricing). Alternatively, there may be severe financial frictions and asymmetries between the incumbent and the rival in terms of their ability to raise external funds.

This discussion also echoes the one in Chapter 2 in relation to conditional (single-product) rebates. Under the presence of a portion of the buyers' demand that is captive to the dominant firm and another portion

⁴³ More precisely, prices play an indirect role when bundling represents a credible commitment to aggressive pricing in case of entry.

⁴⁴ As noted in Section 4.3.4, in the variant of Carlton and Waldman (2002) with network externalities the same effects as bundling can be achieved by engaging in predatory prices in the complementary market. When, instead, the purpose of the incumbent is purely to deter entry, bundling cannot be replaced by predatory pricing.

that is contestable by the rival, the dominant firm engaging in predation may want to offer a discount that is conditional on the buyer's purchases exceeding a given threshold, that is, a discount conditional on the buyer purchasing from the dominant firm not only the captive demand (which corresponds to product *A* in this case) but also the contestable demand (which corresponds to product *B*). The similarity continues in that, in both cases, to identify the actual price that the rival should match so as to capture the contestable demand, one should attribute to the contestable units (to product *B* in this case) the entire discount that the buyer loses when she buys from the rival and does not qualify for the discount.

However, in the case of single-product rebates, designing the discount as conditional has a clear rationale: it allows the dominant firm to target the discount to the contestable portion of the demand, thereby making it less costly to implement predatory prices. In contrast, in the case of multi-product rebates, it is less clear why conditionality is necessary. Since the products involved are distinct, in order to implement a predatory price for product *B*, the dominant firm could in principle simply set a low price for that product, without resorting to commercial strategies that also involve product *A*. In other words, economic theory at the moment does not explain why a dominant firm whose objective is to exclude a rival from market *B* should prefer to resort to bundled rebates instead of predation.⁴⁵

We continue this discussion in Section 4.5, setting out some potential principles that authorities may wish to follow in the assessment of bundled rebates. There, we also discuss in greater detail the possible implementation of a price-cost test, the role of market power in relation to product *B* and the issue of coverage of the bundled rebates (that is, which proportion of the overall product *B* market is covered by the bundled rebates offered by the dominant firm in the product *A* market).

4.3.6 Tying and After-markets

It is quite common for customers to buy a certain product (or original equipment) at a given point in time and then, at a later stage, to buy additional products or services to continue to (or further) benefit from the original purchase. For example, those purchasing a razor will need blades; those buying certain espresso machines will need coffee capsules; owners

⁴⁵ Perhaps specific reasons may justify this preference; for instance, bundled rebates may help 'hide' an exclusionary strategy from competition authorities better than predation.

of a car may need to buy spare parts, as well as maintenance and repair services, if they want to continue using it.

Blades, coffee capsules, spare parts, and maintenance and repair services may be defined as 'after-markets' (or secondary markets) relative to 'primary markets' (razors, espresso machines, cars, respectively).

Shapiro (1995) identifies three defining characteristics for an after-market framework to be applicable: the after-market products are complementary to a primary product and they are all part of the same 'system'; the after-market products (or services) are bought at a later point in time than the purchase of the primary product; and there is some degree of consumer 'lock-in' for example in the form of a switching cost, should the consumer want to move away from the original primary product purchased and select an alternative primary product.

The main issue for antitrust purposes is under what circumstances, if any, a firm may want to monopolise the after-market, and what would be the effects of such a practice. Note also that after-market monopolisation can be achieved in different ways. One could be requirement tying, as mentioned in Section 4.1, which occurs when a company ties the sale of a primary market product to that of an after-market product. For instance, a manufacturer of a machine may impose a contractual obligation to buy all maintenance and after-sales services from it; or it may refuse to honour a warranty unless all the spare parts are bought from it. Very often, though, exclusion of rivals from the after-market can be achieved through practices such as refusal to supply or license (or by setting very high prices for supplying or licensing, which would have the same effect). For instance, the owner of an espresso machine system may invoke patent protection to stop other companies from selling coffee capsules to be used with the Original Equipment Manufacturer's ('OEM's') machine;⁴⁶ or the OEMs of aircraft and engines may refuse to sell spare parts or to provide information necessary for other providers of maintenance and repair services.⁴⁷

⁴⁶ In the EU, there have been a number of cases for instance concerning coffee capsules (or pods); see, for example, *Senseo*, where the Dutch Supreme Court ruled in favour of the independent suppliers (on intellectual property grounds); and *Nespresso*, where the OEM offered commitments in order to address concerns raised by the French Competition Authority.

⁴⁷ See, for example, an article by the Financial Times in early 2016, reporting that some airlines had complained to the European Commission that some OEMs were withholding repair information from qualified providers of third-party maintenance, which the complainants argued restricted competition. See Financial Times, 'Airlines complain to Brussels over parts and maintenance contracts', 23 March 2016.

We shall deal with case-law related to after-market exclusion in Section 4.6, also because such cases represent a sizeable proportion of the case-law on tying. But the reader is advised that we may have also dealt with them in Chapter 5, which discusses practices as refusal to supply or denial of interoperability.

In what follows, we briefly summarise some of the key insights from the academic literature on possible anti-competitive effects in after-markets, which has been spurred by the *Kodak* judgment in the US, which we discuss in Section 4.6.

Before doing so, though, note that from a legal point of view – especially in Europe where an unilateral conduct is abusive only if the firm has a dominant position – one crucial issue is whether a firm which is not dominant in the primary market may be liable for unilateral conduct undertaken on its after-market products or services. This requires, in turn, defining the after-market products or services as separate relevant markets and finding the manufacturer dominant in such market.⁴⁸ In general, one would think that the relevant market should be defined as the system, rather than as primary and after-markets separately: if a firm increases prices in the after-market by a small significant amount, this is unlikely to be profitable if the firm faces competition in the primary market, as consumers – who are assumed to be rational and would consider life-cycle costs when taking decisions – would turn to other primary products.

Only under fairly special circumstances might it be sensible to define separate after-markets for a brand which lacks market power in the primary product, and the discussion below will show some of these circumstances which may lead primary and secondary markets to be somewhat disconnected. For instance, if consumers for some reason have insufficient information about the after-market prices or are behaving naïvely and do not consider the costs they will incur in after-markets; if the value of the particular spare parts or after-sales service was very low with respect to the price of the primary product; or if the probability that a given spare part would break down is so low that it is unlikely to weigh in the original purchase decision.

⁴⁸ For instance, in the examples above, the manufacturer of a branded coffee machine may be found dominant on the technology used in its machine – protected by intellectual property rights – and investigated for monopolising the market for coffee capsules; or an OEM may be found dominant in the market for the production of spare parts of its airplane and accused to monopolise the repair market.

4.3.6.1 The Chicago School view

We can start our discussion from adapting the Chicago School argument reviewed above (see Section 4.3.1) to after-markets. A firm that has a monopoly in the primary market would benefit from lower prices, greater variety or higher quality in the after-market, since these would raise value for the whole system, and a higher price of the primary product will allow to appropriate this value. Hence, an OEM would have no incentive to exclude rivals from the after-market. If it does, it must be because there are some distortions in the after-markets and the tying will then restore efficiency. For instance, suppose that spare parts, maintenance and repair services extend the life-cycle of an equipment. Since the choice of the consumer between replacing the original equipment or repairing it will depend on the relative prices, competition might result in too low prices in the after-market, making the consumers over-demand after-market services, and delay the replacement of the equipment beyond what is efficient. By excluding rivals from the after-market, the argument goes, the OEM would set prices in a more efficient way.

To see the logic, suppose that after-market services are sold at marginal cost because of competition, and the primary product is sold above marginal cost (either because there is a monopolist OEM or because costs of switching from the original equipment when having to replace it would allow for market power). Then, when consumers have to decide between replacing the old equipment or servicing it, their choice will be inefficiently biased towards the latter, and they will use the old equipment longer than it would be optimal.⁴⁹

Similarly, an OEM which wanted to improve the quality of the primary product but which faced fierce competition in the after-market would have to recoup research and development costs by increasing the price of the primary product, which again would distort the choice between replacement and maintenance.⁵⁰

4.3.6.2 Possible rationales for anti-competitive tying in after-markets

Departures from the standard set of assumptions of the Chicago School – for instance by considering insufficient information by consumers, or the impossibility to write contracts which allow a firm to credibly commit to

⁴⁹ See Carlton and Waldman (2010).

⁵⁰ See Elzinga and Mills (2001).

future prices – may however lead to possible anti-competitive motives for excluding rivals from the after-markets.⁵¹

Limited ability to commit to future after-market prices To set the stage, let us refer to a stylised setting inspired by Borenstein et al. (2000), and see under which conditions the Chicago School result would arise. (We provide a more technical discussion in Section 4.4.5.) Suppose that a monopolist produces a durable good, say a machine, which lasts for two periods only. In the second period, the machine depreciates (deteriorates), but a user may buy q units of service which increase the value of continued usage; the price of service, p , will affect the quantity demanded. Assume also that at any moment in time there are N consumers who have bought the machine in the previous period, and who therefore have to decide on how much service to buy, and N consumers who have to buy a new machine. Consumers will continue to take decisions on machine and service purchase period after period. Assume that the firm is a monopolist not only of the machine, but also of its after-market services. In each period, it has to take two decisions, the price P of the machine and the price p of the service. Both prices matter for a consumer at the time of purchasing the machine: the lower the price of the machine, the higher her consumer surplus (just like with most goods); in addition, the lower the expected price of the service, the higher her expected consumer surplus from buying and using the machine over the two periods.

Consider next what happens if the monopolist could write a contract in which it could credibly commit to charge a certain price p of the service in every period. If so, when it starts selling the machine, it will want to set p to equal the marginal cost of providing service (c). In this way, it will maximise the value of the machine to the consumer, and it can then set a price for the machine which allows it to extract all the consumer surplus.⁵²

What we have just set out is nothing but the Chicago School's result: a firm would not have any incentive to set supra-competitive prices in the after-market (service) because it would be able to appropriate any future surplus by suitably choosing the price P of the primary product (the machine).

⁵¹ See also Shapiro (1995), whose classification we follow and whose discussion we draw from.

⁵² Note that by setting p above marginal cost cannot be optimal because of a dead-weight loss: the profit obtained on the after-market would be lower than the surplus lost by the consumer.

An important assumption behind this result is that the seller is able to credibly commit to a certain future after-market price, and could never change that price. But this is a strong assumption, and there will be many circumstances in which such a commitment would be unrealistic. So, let us see now what happens if the monopolist could not make such a commitment. For instance, suppose it has sold the machine to the first cohort of N consumers, and it has to decide whether to charge services at marginal cost or not (and if it has promised to charge $p = c$, it has to decide whether to stick to its promise or not). Its profits will be made of two components: (1) those made by selling after-market services to the installed base (the N consumers who have a machine bought in the previous period); plus (2) all the profits coming from new and future users, whose surplus will be given by the expected value from using the machine over two periods. Sticking to p equal to marginal cost would optimise the second profit component (the problem would be exactly the same as when the monopolist sold the very first machines, and we have seen that setting $p = c$ is optimal), but it would mean not making any profit from the first component. Hence, there would be an incentive to raise the service price p so as to exploit the installed base. Under the assumption that all new and future consumers anticipate that the future price of services is equal to the current price, Borenstein et al. (2000) show that in any period in which there already is an installed base, there is an equilibrium where the monopolist sells services at a price p^* which is strictly above marginal cost in all periods (consumers' expectations – *from that period onwards* – will therefore be correct).⁵³

In sum, a departure from the Chicago School's assumptions to consider a potentially more realistic situation where firms cannot commit to future prices may give rise to a first so-called *limited commitment* theory of harm (due to Borenstein et al., 1995, 2000). Note that the result that after-market prices would be above costs holds not only in monopolistic primary markets. Indeed, the same logic as above also applies to situations

⁵³ Mathematically, by slightly increasing the service price p above cost produces a first-order increase in the profits obtained from the installed customer base. Instead, given that the service price $p = c$ is the one that maximises the flow of profits that the monopolist obtains from selling the machine and the services to the first cohort of customers (when no installed base exists) and to the subsequent ones, such a departure from the optimal price produces a second-order effect on the profits obtained from current and future sales (of the machine and of the services). The intuition in this case is very similar to the one that we discussed in Section 4.3.2, when dealing with mixed bundling as a way to rebalance the prices of two products – one sold under monopolistic condition, the other at marginal cost – so as to increase total profits, and illustrated by Figure 4.1.

where firms imperfectly compete on the primary market (but each of them is the only seller of services for its own machine), and Borenstein et al. (2000) show that services are sold above costs absent the possibility to credibly commit to future after-market prices.⁵⁴ Therefore, even under competition in the primary market, firms will have an incentive to set after-market prices above cost at any period in time. However, as primary products tend to become perfect substitutes, the overall firms' profits tend to zero, with after-market prices above cost but primary product's prices being below cost. Note that after-market rents will be fully competed away only in this extreme case, whereas product differentiation in the primary market would allow for overall (that is, at the system level) positive profits.

Other possible rationales Another possible rationale for anti-competitive tying in after-markets stems from a so-called *surprise theory*, whereby a firm engages in opportunistic behaviour and exploits locked-in buyers. In the example set out above, this is equivalent to saying that after promising the first generation of buyers of its machines that it would set competitive after-market prices, a seller would have an incentive to behave opportunistically and raise such prices – which is precisely what we have just seen it would do.

More generally, suppose that an OEM has in the past promoted an open system, and allowed competition in the after-market for its product, supplying original parts to maintenance and repair service providers, providing them with the information they need, certifying rivals' spare parts and so on. Consumers rationally based their choice of the original equipment on the expectation of low after-market prices, but because of switching costs and high prices of the primary products they may tend to be locked-in with their purchased equipment. If at a point in time the OEM unexpectedly changed its after-market policy and took actions to exclude after-market rivals (for instance, refusing to supply spare parts, instruction manuals, invoking intellectual property rights), it could increase prices of the after-market significantly, and locked-in consumers may have no better choice than to foot the higher bills.

⁵⁴ In fact, the after-market price is even higher under primary market oligopoly than monopoly. Under the latter, a seller is able to appropriate more of the surplus it generates with a low service price through a higher price of the machine; under the former, competition in the primary market constrains the ability to appropriate the surplus through a higher price of the machine. Hence, in this model oligopolists would have an incentive to set a higher after-market price.

Of course, though, this 'surprise' strategy would have consequences, and a firm would have to trade off the higher after-market profits obtained in this way with the expected losses incurred from (a) existing customers prematurely replacing their equipment with a rival one (but switching costs may limit their alternatives) and (b) new customers preferring rival OEM products.

In general, 'surprising' customers with this opportunistic behaviour will then entail a comparison between short-run gains and long-run losses. It will therefore be more likely when the product is mature or declining (several commentators note that 'hold-up' episodes have occurred in sharply declining markets), when it is long-lived (other things being equal, one would expect such opportunism on machines with an expected life of thirty years, say, not on those which are likely to last only five), when the after-market (and its value) is large, when competition in the primary market is weaker and switching costs are more important, when firms have a low discount factor (that is, they value the future relatively little compared to the present), and when the OEM is not active in many product markets (otherwise, it may acquire a bad reputation and be 'punished' by consumers also in other markets).

Shapiro (1995) also notes, however, that consumers will try to protect themselves from this so-called 'installed-base opportunistic behaviour' by demanding contractual clauses such as warranty coverage (whereby at least for some period they would not have to incur the higher costs of after-market services), by renting or leasing equipment rather than buying it, by negotiating long-term service contracts, or a commitment from the supplier to second-sourcing and open systems. These forms of protection may well reduce the extent of the problem but may not always be available.

A further theory of harm of tying in after-markets is based on *costly information* and it concerns situations where buyers face some difficulties in determining the combined price that they will pay for their primary and after-market purchases. (Here again there would be a departure from the 'ideal' market conditions assumed by the Chicago School.) In this case, a supplier may afford charging high prices in the after-market (and would have the incentive to exclude competitors from the after-market) because consumers are not far-sighted and would not consider the overall cost of the system, but just the price of the primary product. This theory of harm would be less likely to apply in the case of expensive products, with large and sophisticated buyers, with repeat buyers or in the presence of third-party sources of information (think of specialist magazines providing estimates of the likely life-cycle costs of cars), when consumers have a

high discount factor (that is, they place relatively large weight on future consumption compared to present one), and in mature markets where information about costs of spare parts and after-sales services and about the probability that a repair is needed is abundant.

Obviously, this theory requires a certain degree of naïvety or short-sightedness on the side of consumers, or very particular circumstances. There is however some empirical evidence showing that consumers may not be as far-sighted as many economists tend to believe.⁵⁵

Still, it may not be necessary for all or most consumers in the market to be far-sighted to have lower after-market prices because of competition on the primary product: a few informed buyers may be sufficient and act as a ‘public good’ for the uninformed ones, unless of course sellers are able to identify and discriminate prices between informed and uninformed buyers (which may happen in markets where buyers and sellers privately negotiate prices).

Shapiro (1995) also notes that – even if indeed most consumers were poorly informed or short-sighted – the expectations of rents in the after-markets will push firms to compete hard in the primary product to appropriate those rents. This is similar to what would happen in any market characterised by switching costs (or network effects, or two-sided markets), where firms set low prices for new (or early, or one group of) buyers, with competition for them dissipating – at least in part – expected rents on older (or successive, or the other group of) buyers. In a market of this type, antitrust intervention aimed at decreasing prices where firms obtain gains might be distortive, as it would deprive firms of the rents that they have sacrificed in the competitive segment of the market. In the case at hand, therefore, high after-market prices might just be the mirror image of the low prices of the original equipment.

Although this argument is unlikely to hold in its extreme version, according to which all after-market rents would be dissipated away by competition in the primary market,⁵⁶ we agree that it should warn against

⁵⁵ For instance, Hausman (1979) analyses the demand for energy-intensive household appliances (for which energy costs represent a very important component of the overall life-cycle expenditures), and finds that consumers tend to give much more weight to the initial capital expenditures than to the subsequent operating expenditures, determining a very high implied discount rate. See Coppi (2007) for references to other empirical works which tend to show customers myopia. See also FCA (2013) for a review by the UK Financial Conduct Authority of possible biases in consumer behaviour, with a focus on financial services.

⁵⁶ See also the discussion above of Borenstein et al. (2000). There is also a small empirical literature which looks at the so-called ‘waterbed effect’, trying to identify what happens to

an interventionist approach in after-markets, also because it points to the fact that anti-competitive damage in the after-market, if any, may be limited in size, as it may at least partially be offset by higher surplus obtained in the purchase of the original equipment.

Finally, tying in after-markets may be based on the *price discrimination (or metering)* rationale discussed in Section 4.2.2. An OEM which is also a monopolist of its after-market product may want to set a low price for its equipment and a higher price in the after-market as a way to discriminate between low-value and high-value consumers. Competition in the after-market would prevent the OEM from following this strategy, thereby giving a possible reason to the OEM to exclude after-market rivals. However, as we have seen in Section 4.2.2, the price discrimination effects of requirement tying are ambiguous, and would typically not justify antitrust intervention.

4.3.6.3 Conclusion

To sum up, there exist a number of theories of harm which may explain why firms may want to charge high after-market prices, and hence to exclude rivals from their after-markets. However, some of these theories may need particular circumstances,⁵⁷ they may involve limited consumer harm, and antitrust action may even be counter-productive, such as in case of an intervention which pushes after-market prices downwards, when firms had already competed away after-market rents through low prices of the primary product.

4.4 Economic Models of Bundling and Tying*

4.4.1 Profitability of Bundling with Elastic Demand Functions*

In this section we will focus on the case in which product *A* and product *B* are not complementary, and generate utility when consumed on a stand-alone basis. Differently from Section 4.3, we now assume that the demand of each product is elastic. We will consider first the case in which the incumbent can extract the entire surplus of consumers

the prices of a component of a system product when regulatory or antitrust intervention forces down prices on the complementary component. For instance, Genakos and Valletti (2011) show that lower termination charges will result in higher prices for mobile subscribers, but the latter increase does not fully compensate for the former decrease.

⁵⁷ See also Coppi (2007), who summarises a number of the factors affecting the likelihood of anti-competitive effects in after-markets.

from the consumption of product *A*, for instance because it can use two-part tariffs (Section 4.4.1.1). We will then analyse the case in which complete surplus extraction is not possible, for instance because firms are constrained to use linear tariffs (Section 4.4.1.2).

4.4.1.1 Perfect rent extraction in a static setting*

Consider a buyer with downward-sloping demand for product *A* and product *B*. Assume that the incumbent firm *I* is the unique supplier of product *A* (with constant marginal cost c_{IA}). Firm *I* produces also product *B*, with constant marginal cost c_{IB} . In market *B* there exist alternative suppliers, producing the same product as the incumbent at marginal cost c_B . Firms can adopt two-part tariffs.

Consider first the case in which firm *I* sells the two products independently. The optimal tariff in market *A* is $t(q_A) = c_{IA}q_A + CS(c_{IA})$: the variable price for product *A* is equal to the incumbent's marginal cost and the fixed fee $T_A = CS(c_{IA})$ allows the incumbent to extract the entire surplus that the buyer enjoys from purchasing at that price. The incumbent's profits from the sales in market *A* amount to $T_A = CS(c_{IA})$. In market *B* competition among alternative producers leads to the tariff $t(q_B) = c_Bq_B$. If the incumbent is less efficient than its rivals, it cannot offer a better deal and it will not sell product *B*. If, instead, the incumbent's marginal cost is lower than that of rivals, it will sell product *B* setting the tariff $t(q_B) = c_{IB}q_B + CS(c_{IB}) - CS(c_B)$. In this case the incumbent's profits from the sales in market *B* amount to $T_B = CS(c_{IB}) - CS(c_B)$: the incumbent appropriates the additional consumer surplus that its more efficient technology generates. The total profits that the incumbent makes under independent sales is $CS(c_{IA}) + \max\{0; CS(c_{IB}) - CS(c_B)\}$.

Consider next the case in which firm *I* sells the bundle. The buyer is willing to choose the bundle from firm *I* if she obtains the same surplus as in the case in which she buys product *B* alone from rivals under the tariff $t(q_B) = c_Bq_B$. Then, in order to persuade the buyer to choose the bundle, firm *I* must leave at least the surplus $CS(c_B)$ to the buyer. As a consequence, the highest surplus that the incumbent can extract through the fixed fee when selling the bundle is $T = CS(c_{IA}) + CS(c_{IB}) - CS(c_B)$. This highlights that bundling gives the incumbent the same profits as in the case of independent sales when the incumbent is more efficient than rivals in the production of product *B*; instead, it gives strictly lower profits when the incumbent is less efficient than rivals as far as product

B is concerned. The decision to bundle the two products is dominated by the decision to sell the two products independently.

The conclusion drawn in the simple case of Section 4.3.1 is still valid: if the incumbent already extracts the entire consumer surplus from the sales of product A stand-alone, and it must leave the buyer at least the same surplus as when she purchases product B stand-alone at a price equal to the rivals' marginal cost, then there is no additional surplus that it can appropriate by bundling the two products.

4.4.1.2 Imperfect rent extraction in a static setting*

Assume now that firms are constrained to use linear tariffs and, let us focus for simplicity on the case in which the incumbent is as efficient as rivals in the production of product B : $c_{IB} = c_B$. In this section, we denote with c_A the marginal cost of product A .

Consider first the case in which firm I sells the two products independently. It will set the monopoly price p_A^m in market A . In market B the equilibrium price is c_B . The incumbent's total profits amount to the monopoly profits π_A^m from product A .

Consider now the case in which firm I decides to bundle the two products (whose prices are denoted by p_A^b and p_B^b). Assume that it still sells product A as a stand-alone product at the price p_A^m . Consumers can also obtain product B stand-alone at the competitive price c_B . The optimal prices in the bundle are the ones solving the following problem:

$$\max_{p_A^b, p_B^b} (p_A^b - c_A)D_A(p_A^b) + (p_B^b - c_B)D_B(p_B^b) \quad (4.4)$$

$$s.t. \quad CS_A(p_A^b) + CS_B(p_B^b) \geq CS_A(p_A^m) + CS_B(c_B) \quad (4.5)$$

Greenlee et al. (2008) show that the prices (p_A^{b*}, p_B^{b*}) that solve this problem are such that $c_i < p_i^{b*} < p_i^m$ with $i = A, B$. The price p_i^m is the monopoly price in market i , that is, the price that solves: $D_i(p_i) + (p_i - c_i)D_i'(p_i) = 0$. In other words, the optimal prices are below the monopoly price and above the marginal cost of each market. Moreover, the incumbent's total profits increase as compared to independent sales: $\pi_A^* + \pi_B^* > \pi_A^m$.

Proof. The above optimisation problem can be written as:

$$\max_{p_A^b, p_B^b} \left\{ (p_A^b - c_A)D_A(p_A^b) + (p_B^b - c_B)D_B(p_B^b) + \right.$$

$$\lambda \left[\int_{p_A^b}^{p_A^m} D_A(s) ds - \int_{c_B}^{p_B^b} D_B(s) ds \right] \quad (4.6)$$

The first-order conditions are:

$$D_A(p_A^b) + (p_A^b - c_A)D_A'(p_A^b) - \lambda D_A(p_A^b) = 0 \quad (4.7)$$

$$D_B(p_B^b) + (p_B^b - c_B)D_B'(p_B^b) - \lambda D_B(p_B^b) = 0 \quad (4.8)$$

$$\lambda \left[\int_{p_A^b}^{p_A^m} D_A(s) ds - \int_{c_B}^{p_B^b} D_B(s) ds \right] = 0 \quad (4.9)$$

First we show that the constraint binds at the optimal solution. Imagine that this is not the case, so that $\lambda = 0$. Then, equations (4.7) and (4.8) are satisfied by $p_A^b = p_A^m$ and $p_B^b = p_B^m$ and the constraint $CS_A(p_A^b) + CS_B(p_B^b) \geq CS_A(p_A^m) + CS_B(c_B)$ is not satisfied. This leads to the conclusion that $\lambda > 0$.

From $\lambda > 0$ it follows that the prices (p_A^{b*}, p_B^{b*}) that solve equations (4.7) and (4.8) are strictly lower than the monopoly prices (p_A^m, p_B^m) , where p_i^m (with $i = A, B$) is such that $D_i(p_i^m) + (p_i^m - c_i)D_i'(p_i^m) = 0$. Then, for the constraint to be binding with $p_i^{b*} < p_i^m$, it must be that $p_B^{b*} > c_B$.

From equation (4.8), $p_B^{b*} > c_B$ implies that $\lambda < 1$, so that also $p_A^{b*} > c_A$.

Finally, since the prices (p_A^m, c_B) satisfy the constraint and could have been chosen, then the optimal prices (p_A^{b*}, p_B^{b*}) strictly increase total profits of the incumbent. ■

4.4.2 Carlton and Waldman (2012)*: Imperfect Rent Extraction from Future Investments

Let us assume that there are two complementary products, the primary one, denoted by P , and the secondary one, denoted by S . There exists a unit mass of identical consumers that value U the system product (that is, the combination of products P and S) for every period when they consume it, whereas buying either product alone would give zero utility. The incumbent is the unique producer of the primary product, while it faces competition in the production of the secondary product. The incumbent's marginal cost of production is c for the primary product and c_I for the secondary product. The rival firm is more efficient than the incumbent in the production of the secondary product: $c_R = 0 <$

c_I .⁵⁸ The secondary product is perfectly homogeneous: there are no differences between the incumbent's and the rival's secondary products.

The timing of the game is the following. There are two periods, period 1 and period 2. Both the primary and the secondary products are durable, so that a unit purchased in period 1 can be used in both periods.

At the beginning of period 1, the incumbent chooses whether to sell the two products in a bundle or to sell them separately. The bundling decision is irreversible. Moreover once the consumer has purchased the bundle from the incumbent, she cannot add the rival's secondary product to the system (this may be due to the fact that the bundling decision also involves denying the rival of the information necessary to make its secondary product compatible with the incumbent's primary product, or to the consumer having a strong preference for single-homing, perhaps driven by significant switching costs or lack of storage space, or other reason). After the incumbent makes its decision on whether to bundle, firms simultaneously choose prices for their products and consumers choose which products to purchase.

In period 2, firms first simultaneously choose whether to upgrade their secondary products, which entails a fixed cost of $Z > 0$. When an upgraded secondary product is incorporated, the utility generated by the system is increased by λ . The production costs of the upgraded secondary products are the same as the ones of the non-upgraded products (that is, c_I for the incumbent and $c_R = 0$ for the rival). Then firms simultaneously choose prices for the upgraded secondary products (if any) and consumers choose whether and from which firm to buy.

The restrictions on the parameters are as follows:

$$2U - c - c_I > 0 \quad (4.10)$$

$$\lambda > Z + c_I \quad (4.11)$$

$$c_I > Z \quad (4.12)$$

Condition 4.10 ensures that production of the basic system (that is, without upgrading) is efficient. Condition 4.11 ensures that upgrading is efficient. Finally, condition 4.12 ensures that the rival firm has an incentive to invest in upgrading even when the incumbent also invests.

⁵⁸ Equivalently, we could assume that costs are identical but the rival's product is of higher quality ($v_R > v_I$).

For simplicity, the discount factor is assumed to be $\delta = 1$. Moreover, there is no scrap value and no second-hand market for the secondary product in period 2. Finally, an important assumption is that in period 1 consumers are not willing to pay for the surplus generated by an investment that has not materialised yet in period 1; moreover, in period 1 the incumbent is assumed not to be able to set prices contingent on the realisation of upgrading in period 2 (for instance, it cannot set an overcharge for the primary product due in period 2 if upgrading occurs). These assumptions imply that in period 1 the incumbent cannot extract from consumers the additional surplus that they may enjoy in period 2 from the upgrading of the secondary product.

As usual we solve the model by backward induction. We consider first the case in which the incumbent decides to bundle the primary and secondary products in period 1 (Section 4.4.2.1); then we will consider the case in which it sells the two products separately (Section 4.4.2.2); finally, we will study the bundling decision (Section 4.4.2.3).

4.4.2.1 Bundling*

If the incumbent decides to bundle product P and product S , the rival can sell neither its secondary product in period 1 nor its upgraded secondary product in period 2. Then the rival does not invest in upgrading in period 2. In period 1, the incumbent sets a price $\tilde{p}^1 = 2U$ for the bundle, it invests in upgrading at the beginning of period 2 and it produces and sells the upgraded secondary product in period 2 at a price $p_{I,S}^2 = \lambda$. (Assumption 4.11 ensures that upgrading is profitable for the incumbent.) In this case, the total profits of the incumbent amount to $\pi_I^{1+2,b} = 2U - c - c_I + \lambda - c_I - Z$.

4.4.2.2 No bundling*

Let us next turn to the case where the incumbent does not bundle. First, we analyse the investment decision in period 2.

Suppose the incumbent does not invest. Then, if the rival invests, it will set a price $p_{R,S}^2 = \lambda$ and consumers will buy the upgraded secondary product. (Note that when the incumbent sells the two products independently, in period 2 consumers can substitute any secondary product purchased in period 1 with any upgraded product purchased in period 2.) Investing is profitable because $\lambda - Z > 0$ by assumption 4.11. In this case the incumbent's profit in period 2 is zero.

Suppose now that the incumbent invests. If the rival also invests, then they compete in period 2 with homogeneous upgraded secondary products. Since the rival is more efficient than the incumbent, it will sell the upgraded product setting a price $p_{R,s}^2 = c_I$. By assumption 4.12, $c_I - Z > 0$ and the investment is profitable. In this case the incumbent does not sell the secondary product and does not cover the investment costs Z : $\pi_I^2 = -Z$.

Since the rival always invests (irrespective of the incumbent's decision), then the incumbent will decide not to invest at the beginning of period 2.

Let us move now to period 1. By assumption, in period 1 firms can at most extract the surplus generated in the two periods by the basic system (primary product plus non-upgraded secondary product). Since the rival is more efficient than the incumbent in the production of the secondary product, the rival will supply the consumers. However, there are multiple equilibria in the price game: the equilibrium price in the secondary market is $p_{I,s}^1 = p_{R,s}^1 = (1 - \alpha)c_I$, with $\alpha \in [0, 1]$, and the incumbent's price in the primary market is $p_{I,p}^1 = 2U - (1 - \alpha)c_I$. Note that the rival appropriates its entire efficiency advantage $c_I - c_R = c_I$ when $\alpha = 0$. Instead, when $\alpha > 0$ the rival's margin in the secondary market is (at least partially) squeezed, which allows the incumbent to set a higher price for the primary product, thereby appropriating a share α of the increase in surplus generated by the activity of the more efficient rival. For simplicity we assume here that $\alpha = 1/2$, but this assumption does not affect the result of the model. Then, when the two products are sold independently, the incumbent sells the primary product in period 1; and the rival sells the secondary product in period 1, it invests in period 2 and it then sells the upgraded secondary product in period 2. Total profits for the incumbent amount to $\pi_I^{1+2,nob} = 2U - c - c_I/2 + 0$.

4.4.2.3 Bundling decision*

By comparing the incumbent's total profits with and without bundling, one can note that bundling harms the incumbent, because it prevents the incumbent from appropriating some of the surplus produced by the activity of the more efficient rival in the secondary market in period 1. However, it also benefits the incumbent, because by bundling it is the incumbent that invests in upgrading and appropriates the benefits of the investment. When the additional profits generated by upgrading are

large enough (that is, where λ is large enough), the incumbent finds it more profitable to bundle the products in period 1:

$$\begin{aligned}\pi_I^{1+2,b} &= 2U - c - c_I + \lambda - c_I - Z > 2U - c - c_I/2 \\ &= \pi_I^{1+2,nob} \Leftrightarrow \lambda > Z + 3c_I/2\end{aligned}\quad (4.13)$$

Note that when it occurs, bundling is welfare-detrimental because it forecloses the more efficient producer of the secondary product from that market. Note also that bundling is the only way for the incumbent to discourage the rival's investment.⁵⁹

4.4.3 Carlton and Waldman (2002)*: The Model with Fixed Entry Costs

We now provide a formal treatment of Carlton and Waldman (2002). We will first analyse the variant of the model where scale economies arise on the supply-side due to the presence of fixed entry costs. We will then analyse the variant of the model with network externalities. Let us assume that there are two complementary products, the primary one, denoted by P , and the secondary one, denoted by S . Consumers value U the system product, whereas buying either product alone would give zero utility.

The entrant, firm E , is as efficient as the incumbent, firm I , in the production of the primary product – their common marginal cost is c – but it is more efficient than the incumbent in the production of the secondary product: $c_E = 0 < c_I$.

In the primary market, the incumbent enjoys a safe monopoly position in the first period, while entry is possible in the second period; in the secondary market, entry is possible in both periods. Entry requires to pay a fixed and sunk cost f_p for the primary market and f_s for the secondary market. The restrictions on entry costs are the following:

$$c_I < f_p + f_s < c_I + c_I/2 \quad (4.14)$$

$$f_p < c_I/2 \quad (4.15)$$

⁵⁹ If the incumbent set a very low price for the secondary product in period 1 (say a price equal to zero), it could discourage the rival from selling the secondary product in period 1, but it would not discourage second-period investment.

The role of these restrictions will become clear as the solution of the game proceeds. Firm I , being incumbent, has already paid the entry costs before the game starts.

The game is as follows. At period 0, firm I decides whether it wants to sell the two products as a bundle, or separately. The decision to bundle is irreversible. At period 1, (i) firm E decides whether it wants to enter the secondary market, and accordingly pay f_s , or not; (ii) price choices are made by active firms; (iii) buyers decide, and (iv) transactions are made and profits realised. At period 2, (i) firm E decides whether it wants to enter the primary product market, and accordingly pay f_p , or not (and if it has not entered the secondary market in period 1, it has another chance to do so); (ii) price choices are made by the two firms; (iii) buyers decide, and (iv) transactions are made and profits realised. We assume no discounting between the two periods.

As usual, we solve the game by backward induction. We consider first the case where the incumbent chooses to sell the two products separately (Section 4.4.3.1), and then we will move to the case of bundling (Section 4.4.3.2).

4.4.3.1 No bundling*

Period 2

Let us begin from period 2. We need to distinguish two cases, one where firm E entered the secondary market in period 1; the other where it did not.

Firm E entered the secondary market in period 1

In this case, in period 2 firm E faces two alternatives: it can either also enter the primary market or it can stay out of it.

Case I: Firm E does not enter the primary market in period 2 The incumbent chooses the price of the primary product free of competitive pressure; however, as we already discussed in Section 4.3 the price that the incumbent can set for the primary product is linked to the equilibrium price of the secondary product. The lower the latter price, the higher the price $p_{I,p}^*$. Since firm E is more efficient than the incumbent in the secondary market, there are multiple equilibria in the price game: the equilibrium price in S is $p_{I,s}^* = p_{E,s}^* = (1 - \alpha)c_I$,

with $\alpha \in (0, 1)$,⁶⁰ and with the entire market being covered by firm E . The incumbent chooses the price $p_{I,p}^* = U - (1 - \alpha)c_I$ in the primary market.

Note that firm E appropriates its efficiency advantage $c_I - c_E = c_I$ entirely only when $\alpha = 0$, and the equilibrium price in the secondary market is c_I . Instead, when $\alpha > 0$ the entrant's margin in the secondary market is squeezed, which allows the incumbent to set a higher price for the primary product, thereby extracting a share α of the increase in buyers' surplus generated by the activity of the more efficient rival.⁶¹ Without loss of generality, the assumption made in Carlton and Waldman (2002) is that $\alpha = 1/2$, that is, $p_{I,s}^* = p_{E,s}^* = c_I/2$ and $p_{I,p}^* = U - c_I/2$. Consequently, the agents' payoffs in the second period are:⁶²

$$\pi_{I,2} = U - c_I/2 - c; \quad \pi_{E,2} = c_I/2; \quad CS_2 = 0 \quad (4.16)$$

Case II: Firm E does enter the primary market in period 2 In this case, Bertrand competition leads to $p_{E,p}^* = p_{I,p}^* = c$ in the primary market and to $p_{E,s}^* = p_{I,s}^* = c_I$ in the secondary market (with the entrant selling in the secondary market). Note that, due to firm E 's presence in the primary market, now the incumbent cannot increase the price of product P above c . Then, the entrant's margin is not squeezed and firm E can fully appropriate its efficiency advantage in the secondary market. The agents' payoffs in the second period, gross of the entry cost, are:

$$\pi_{I,2} = 0; \quad \pi_{E,2} = c_I; \quad CS_2 = U - c - c_I \quad (4.17)$$

Comparing equation (4.16) to equation (4.17), one can notice that firm E benefits from entry in the primary market even though symmetric Bertrand competition in that market drives profits down to zero. The reason why entry in P is beneficial is that it allows firm E to avoid margin squeeze in the secondary market, thereby increasing the margin that it makes on product S from $c_I/2$ to c_I .

⁶⁰ We will explain later the reason why α cannot take the value of 0 or 1.

⁶¹ When $\alpha > 0$ the incumbent is offering a below-cost price in the secondary market, but earns more in the primary market. If buyers should make a mistake and buy from the incumbent, it would earn $U - c_I - c$, which is the same amount it earns if it offers a price equal to its marginal cost in the secondary market.

⁶² To rationalise this assumption, assume a slightly different version of the price stage of the game in each period. Suppose that with probability 1/2, it is the incumbent that chooses prices first, and with a probability 1/2 it is the entrant that chooses first. In the former

Then, firm E decides to enter the primary market in the second period if such profit increase is larger than the entry cost:

$$c_I - \frac{c_I}{2} > f_p. \quad (4.18)$$

The inequality above is satisfied by assumption (4.15). To sum up, if firm E entered the secondary market in period 1, then it decides to enter the primary market in the second period.⁶³

Firm E did not enter the secondary market in period 1

In this case firm E faces four alternatives in the second period. (i) It can decide to stay out of both markets, in which case its profits are zero. (ii) It can enter the primary market only. In this case Bertrand competition between symmetric firms determines $p_{I,p}^* = p_{E,p}^* = c$, and firm E is unable to cover the entry cost f_p . (iii) It can enter the secondary market only. As showed above, in this case, firm E makes gross profits $c_I/2$ which is not large enough to cover the entry cost f_s .⁶⁴ (iv) It can enter both markets. In this case, as shown above, the gross profit of firm E is c_I which, by assumption (4.14) is insufficient to cover the total entry cost $f_p + f_s$. In other words, the rents that firm E makes in the second period in the two markets are insufficient to make entry profitable.

To sum up, if firm E did not enter the secondary market in period 1, no entry is the most profitable alternative in period 2.

Period 1

Let us move to the first period. The entrant anticipates that, if it does not enter the secondary market, then no entry will take place in the following period. Total profits in this case are zero. If, instead, it enters the secondary market, its total profits, net of the entry costs are:

$$\pi_E^{TOT} = \frac{c_I}{2} + c_I - f_s - f_p \quad (4.19)$$

case, at equilibrium I would set $p_{I,s} = c_E = 0$, $p_{I,p} = U$ and, in the following period, E would set $p_{E,s} = 0$ resulting in profits $\pi_{I,2} = U - c$ and $\pi_{E,2} = 0$. In the latter case, E would set $p_{E,s} = c_I$, and subsequently I would choose $p_{I,s} = c_I$ and $p_{I,p} = U - c_I$. Profits would be $\pi_{I,2} = U - c - c_I$ and $\pi_{E,2} = c_I$. *Ex ante*, expected profits would therefore be: $\pi_{I,2} = U - c - c_I/2$, $\pi_{E,2} = c_I/2$.

⁶³ This analysis explains why the value $\alpha = 0$ needs to be excluded. In that case the entrant would appropriate its entire efficiency advantage even without entering market P . Then, it would have no reason to pay the fixed cost f_p .

⁶⁴ The assumptions $f_p + f_s > c_I$ and $f_p < c_I/2$ imply that $f_s > c_I/2$.

In fact firm E collects rents $c_I/2$ in the first period and c_I in the second period when it also enters market P . By assumption (4.14) such total rents are sufficient to cover the total entry costs. Note that firm E 's ability to collect rents in the secondary market in period 1 is key to make the overall entry decision profitable: as shown above, the rents collected in the second period only are insufficient to make entry in both markets viable.⁶⁵

In sum, if there is no bundling, the unique continuation equilibrium of the game is such that entry will take place in both markets.

4.4.3.2 Bundling*

Let us now study the case where there is bundling, starting from period 2.

Period 2

Also, in this case, we need distinguish the situation where firm E entered the secondary market in period 1 from the situation where it did not.

Firm E entered the secondary market in period 1

Since the incumbent sells a bundle of the two products, firm E will not be able to sell in the second period if it decides not to enter the primary market. In this case, $\pi_{E,2} = 0$. Instead, as shown above, the entrant will be able to compete with the incumbent if it enters the primary market in the second period, setting a price c for the primary product and c_I for the secondary product. Then, its second period profit, net of the entry cost, is $c_I - f_p > 0$ by assumption (4.15). Then, if firm E entered the secondary market in period 1, entry in the primary market will follow in period 2.

Firm E did not enter the secondary market in period 1

Similarly to the case where there is no bundling, no entry is the most profitable strategy in period 2. Indeed, the rents collected in the second period only if firm E enters both markets are insufficient to cover the total entry costs: $c_I < f_p + f_s$. If the entrant enters only the secondary market, it will not be able to sell, due to the bundling choice of the incumbent. Hence, it will not cover the entry cost f_s . The same

⁶⁵ This remark explains why the value $\alpha = 1$ must be excluded. In such a case, when firm E sells only in the secondary market, its margin would be entirely squeezed. Hence, it

outcome occurs if it only enters the primary market: it will make no sales and will be unable to cover the entry cost f_p .

Period 1

When the incumbent chose to bundle the two products, firm E does not sell the secondary product in the first period if it enters such a market. Since the rents c_I collected in the second period, when entry in the primary market takes place, are insufficient to cover the total entry costs $f_p + f_s$, then firm E prefers to stay out of the secondary market in period 1, and from the primary market in period 2.

4.4.3.3 Period 0: Bundling decision*

If the incumbent chooses bundling, then firm E will not be able to collect rents in the secondary market in period 1, which makes overall entry (in market S in period 1 and in market P in period 2) unprofitable. This allows the incumbent to preserve its monopoly position in the primary market and to charge the price U in both periods, thereby making total profits equal to:

$$\pi_I^{TOT(bundling)} = U - c - c_I + U - c - c_I \quad (4.20)$$

If, instead, the incumbent chooses not to bundle, entry will occur in the secondary market in period 1 and in the primary market in period 2. The incumbent's total profits are:

$$\pi_I^{TOT(no\ bundling)} = U - c - \frac{c_I}{2} + 0 \quad (4.21)$$

Note that, absent bundling, in period 1 the incumbent earns more than in the case of bundling: $U - c - c_I/2 > U - c - c_I$. The reason is that, absent bundling, entry occurs in the secondary market in period 1, and the incumbent appropriates half of the surplus created by the activity of the more efficient rival. Then, if there was no threat of entry in the primary market, the incumbent would have no incentive to engage in bundling. However, absent bundling, entry in the secondary market will be followed by entry in the primary market, which will prevent the incumbent from extracting rents in the second period. Instead, with bundling entry in P is prevented and firm I will earn $U - c - c_I > 0$ in period 2. Then, by

would collect zero rents in the first period and entry would never be profitable even absent bundling.

using bundling, the incumbent loses $c_I/2$ in the first period, but gains $U - c - c_I$ in the second period. Bundling is profitable iff

$$U - c - c_I > c_I/2 \iff U - c > 3c_I/2 \quad (4.22)$$

Of course bundling reduces total welfare because the secondary product is produced by the inefficient incumbent.

Discussion

The variant of Carlton and Waldman (2002) with fixed entry costs is close to the variant of the predation model, discussed in Chapter 1 (see, specifically, Section 1.2.3.2 and technical Section 1.3.2.6), where there are two markets related by scope economies, for instance due to the presence of fixed costs common to the production of the two products. The incumbent is already present in all the markets; by contrast, the rival can enter/expand in one market only in period 1, while entry/expansion is possible also in the other market in period 2. Since second-period total demand is insufficient for the rival to reach efficient scale, the incumbent may predate in the market which opens first, to secure such crucial market in period 1, and preserve its dominant position across all the markets where it is active.

There are some important differences, though, between the two models. First, Carlton and Waldman (2002) assume that the first-period entry cost has to be paid *before* price offers are made. As a consequence, in order to deter entry, tying must be decided before entry materialises and must be irreversible. Hence, the incumbent must enjoy a first-mover advantage which is not necessary in the predation model. The difference in timing also explains why, in Carlton and Waldman (2002), entry cannot be deterred simply by using aggressive pricing of the complementary product in period 1, as it happens in the predation model. Second, and more importantly, Carlton and Waldman's (2002) is essentially a model of entry deterrence, where scale economies are due to fixed and sunk entry costs, whereas the model developed in Chapter 1 predatory pricing can not only deter entry but also limit expansion of an existing competitor (as long as scale matters for efficiency).

4.4.4 Carlton and Waldman (2002)*: The Model with Network externalities

In this section, we present a simplified version of the variant of Carlton and Waldman (2002) where scale economies arise on the demand-side

due to the existence of network externalities in the secondary market. We keep the same assumptions and notation as in the previous section. In particular, we assume that the incumbent and the entrant are equally efficient in producing the primary product, whose marginal cost is c , while the entrant is more efficient in producing the secondary product: $c_I > c_E = 0$. Moreover, the entrant can enter the secondary market in any period, but it can enter the primary market only in the second period. The incumbent does not have any fixed costs to pay. The entrant has to pay the fixed cost f_p to start operations in the primary market whereas, unlike the previous variant of the model, there is no fixed cost to operate in the secondary market. The fixed cost f_p is assumed to satisfy the following conditions:

$$0 < f_p < \frac{N(U - c)}{2} \quad (4.23)$$

We assume there are N consumers who can make purchases in the first period, and N in the second period. For simplicity, all of them start to consume only at the end of period 2.⁶⁶

Consumption of the secondary product generates network externalities. We model network effects in a cruder (but simpler) way than in Carlton and Waldman (2002): only if it is used by a sufficient number of consumers, the secondary product, once joined to the primary one, does it generate utility. In other words, the system given by the primary product and the incumbent's (respectively, entrant's) secondary product will give utility $v(n) = 0$ if the secondary product is bought by $n < 2N$ consumers, and utility $v(n) = U$ if the secondary product is bought by $n \geq 2N$ consumers.

Finally, we follow Carlton and Waldman (2002) (and most of the literature on network industries – see, for example, Katz and Shapiro, 1986) by assuming that consumers of a certain cohort choose as if they were able to coordinate. In formal terms, we restrict attention to coalition-proof perfect Nash equilibria.

The game is as follows. At period 0, firm I decides whether it wants to sell products P and S as a bundle, or separately. At period 1, (i) price choices are made by the two firms (recall that firm E can sell the secondary product, because there are no entry costs to enter market S);

⁶⁶ This could also be rationalised by saying that the good is going to be sold only in periods 1 and 2, but then is consumed for an infinite number of periods in the future. As long as the discount factor is large enough, period 1 consumption will become irrelevant.

prices cannot be negative; (ii) first-period buyers decide. At period 2, (i) firm E decides whether it wants to enter the primary market, and accordingly pays f_p , or not; (ii) price choices are made by the two firms; (iii) buyers decide; and (iv) all consumption takes place and profits are realised.

We solve the game by backward induction starting from the case where the incumbent, in period 0, chooses to sell the two products independently.

4.4.4.1 No bundling*

Period 2

In period 2 firm E decides whether to enter the primary market. Its choice depends on the decision taken by the first cohort of buyers in period 1.

Suppose that in the first period all consumers bought from firm E . If firm E enters the primary market, there will be symmetric Bertrand competition in that market, resulting in the equilibrium price $p_p^* = c$ and either of the firms producing the product. Given the choice of the first cohort of consumers, the incumbent's secondary product, joined with the primary one, has too low quality to generate utility to the second cohort of consumers. Then, firm E will sell the secondary product to the second cohort of customers setting the price $p_{Es}^* = U - c$. The second-period payoffs will be:

$$\pi_{E2} = N(U - c); \quad \pi_{I2} = 0; \quad CS_2 = 0. \quad (4.24)$$

If, instead, firm E decides not to enter the primary market, then there will be a continuum of Nash equilibria with firm I selling the primary product at a price $p_{Ip}^* \in [c, U]$ and firm E selling the secondary product at a price $p_{Es}^* \in [0, U - c]$. By assuming that the two firms bargain on prices and that the surplus is being divided equally, we have the following payoffs:

$$\pi_{E2} = \frac{N(U - c)}{2}; \quad \pi_{I2} = \frac{N(U - c)}{2}; \quad CS_2 = 0. \quad (4.25)$$

Note that, by entering the primary market, firm E manages to appropriate the entire surplus generated by its superior secondary product rather than half of it as it happens when it does not enter the primary market. Therefore, firm E will find it profitable to enter if the

additional surplus that it extracts exceeds the entry cost:

$$\frac{N(U - c)}{2} > f_p. \quad (4.26)$$

This inequality is satisfied by assumption (4.23). Hence, if the first cohort of consumers decides to buy from firm E , then entry in the primary market will take place and the second cohort of consumers will buy from firm E .

Suppose now that in the first period all consumers bought the secondary product from the incumbent. Given the choice of the first cohort of consumers, the entrant's secondary product joined with the primary one has too low quality to generate utility to the second cohort of consumers. Hence, firm E will not be able to sell in the secondary product in period 2 irrespective of whether or not it enters the primary market. Moreover, due to symmetric Bertrand competition, if it enters the primary market, the equilibrium price will go down to marginal costs. Hence, it will not be able to cover the entry cost f_p . Then, when the first cohort of consumers chose the incumbent, firm E will find it profitable to stay out of the secondary market in period 2. The incumbent monopolises the two markets extracting the entire surplus of consumers. Payoffs are as follows:

$$\pi_{E2} = 0; \quad \pi_{I2} = N(U - c - c_1); \quad CS_2 = 0. \quad (4.27)$$

Period 1

Let us move to the first period. The first cohort of consumers anticipates that the supplier that they choose will also serve the second cohort of consumers. Therefore, their net payoff from buying the incumbent's secondary product is $U - p_{Ip}^1 - p_{Is}^1$ whereas the net payoff from buying the entrant's secondary product is $U - p_{Ip}^1 - p_{Es}^1$. It follows that, for a given p_{Ip}^1 , consumers will choose the cheaper supplier of the secondary product.

Suppose that the incumbent offers p_{Ip}^1 and p_{Is}^1 while the entrant slightly undercuts p_{Is}^1 and captures the first cohort of consumers. In this case, firm E would enter the primary market in period 2 and would capture also the second cohort of consumers, thereby making total profits equal to:

$$\pi_E^{TOT} = N(p_{Is}^1) + N(U - c) - f_p. \quad (4.28)$$

Then, by assumption (4.23), there is no feasible price $p_{I_s}^1 \geq 0$ that the incumbent can set so as to discourage firm E 's undercutting and to secure the first cohort of customers.⁶⁷ There will be a continuum of Nash equilibria where the entrant sells the secondary product to the first cohort of consumers at the price $p_{E_s}^* = (1 - \alpha)c_I$ and the incumbent sells the primary product at the price $p_{I_p}^* = U - c_I + \alpha c_I$ with $\alpha \in [0, 1]$. These equilibria differ in terms of the extent to which the incumbent manages to extract the rival's surplus through the price of the primary product. Assuming $\alpha = 1/2$, one obtains the following total payoffs:

$$\pi_E^{TOT} = N\left(\frac{c_I}{2}\right) + N(U - c) - f_p; \quad \pi_I^{TOT} = N\left(U - c - \frac{c_I}{2}\right); \quad CS^{TOT} = 0. \quad (4.29)$$

4.4.4.2 Bundling*

Period 2

If there has been bundling in the first period, all first-period consumers could only have bought the secondary product from the incumbent. Therefore, in the second period nobody will want to buy the secondary product from the entrant. Further, there is no point for the entrant to enter the primary market, so the incumbent monopolises both markets and make profits $\pi_{I2} = N(U - c - c_I)$.

Period 1

Clearly, bundling allows the incumbent to be the only seller of the secondary product in the first period, so that total payoffs are:

$$\pi_E^{TOT} = 0; \quad \pi_I^{TOT} = 2N(U - c - c_I); \quad CS^{TOT} = 0. \quad (4.30)$$

4.4.4.3 Period 0: Bundling decision*

At period 0 the incumbent faces the following trade-off: by choosing to bundle, it manages to keep the entrant out of the primary market and to obtain second-period monopoly profits $N(U - c - c_I)$ rather than zero second-period profits as in the case where it chooses to sell the two products independently and entry occurs. However, this occurs at the

⁶⁷ Instead, in Carlton and Waldman (2002), the incumbent may be able to secure the first cohort of consumers through an aggressive (non-negative) price of the secondary product, thereby securing also the second cohort of consumers and discouraging entry

cost of forgoing the rents that it can extract in the first period from the entrant when it does not choose to bundle and firm E secures the first cohort of customers. Indeed, by choosing to bundle the incumbent obtains a profit $N(U - c - c_I)$ in the first period which is lower than the profit $N(U - c - c_I/2)$ that it makes in the first period absent bundling. Bundling is the optimal choice for the incumbent if:

$$N(U - c - c_I) > N\left(\frac{c_I}{2}\right) \iff U - c > \frac{3c_I}{2}$$

Finally, concerning welfare, without bundling total welfare amounts to:

$$W^{no\ bundling} = 2N(U - c) - f_p,$$

whereas under bundling total welfare amounts to:

$$W^{bundling} = 2N(U - c - c_I).$$

Therefore, bundling is welfare-detrimental if:

$$f_p < 2Nc_I.$$

This is the usual ‘excess entry’ argument (see Mankiw and Whinston, 1986) according to which an entrant might have more incentives to enter into an industry than is socially optimal. When the fixed cost of entry is large enough, entry deterrence prevents a duplication of investments from arising, and this turns out to be beneficial from the point of view of welfare.

Discussion

The network externality variant of Carlton and Waldman (2002) is close to the variant of the predation model with network externalities discussed in Chapter 1 (see specifically Section 1.2.3.2 and technical Section 1.3.2.7). In particular, the assumption made in this second variant of Carlton and Waldman (2002) that in the secondary market entry costs are zero implies that, differently from the variant with supply-side scale economies: (i) the incumbent does not need to rely on a first-mover advantage to exclude; and (ii) not only tying but also low pricing for the secondary product in period 1 may be an effective

in the primary market. In other words, in Carlton and Waldman (2002) *virtual* tying (through aggressive pricing) may allow the incumbent to reproduce the same outcome and profits as under *actual* tying. The different formalisation of the network externalities, which here takes an extreme form, is the main driver of this difference.

exclusionary practice. (As discussed above, in the simplified version of the model we have presented here, this possibility does not arise because of the extreme way we have adopted to model network externalities. In a more general framework, aggressive pricing of the secondary product in the first period can reproduce the same outcome as bundling, as shown by Carlton and Waldman 2002.) This makes Carlton and Waldman (2002) even closer to the predation model presented in Chapter 1.

The main difference between the two models is, also in this case, that in Carlton and Waldman (2002) (virtual) tying is profitable for the incumbent only if it leads to entry deterrence in the primary market, whereas in the model in Chapter 1 predatory pricing is profitable also when it marginalises an existing rival. More precisely, in Carlton and Waldman (2002), the key reason why the incumbent may manage to price more aggressively than the rival in the first period is that the rival has not paid the entry costs to operate in the primary market yet. Hence, if the incumbent secures the first cohort of customers, then it will be able to monopolise the system market in the second period. Instead, if the rival secures the first cohort, then it will always have to compete with the incumbent and will extract only duopoly rents in the second period. Instead, in the predation model the key reason why the incumbent manages to price more aggressively than the rival is that it relies on a larger installed base of customers. This makes the quality of the incumbent's network when supported by a single cohort of buyers larger than the rival's and enables the incumbent to extract more rents than the rival from the second cohort of customers, once the first cohort has been secured.

Finally, note that due to the absence of the entry costs in the secondary market, in the network externality variant of Carlton and Waldman (2002) the tying decision need not be irreversible, differently from what happens in the fixed costs variant of the model. The reason is that, in that variant, protection of the incumbent's position in the primary market requires the rival's *entry* decision in the secondary market to be discouraged, that is, its decision to pay the fixed entry cost to operate in that market in the first period. Instead, in the network externality variant of the model, where there are no entry costs in the secondary market, exclusion requires the rival to be prevented from *selling* in the secondary market in the first period.

4.4.5 Tying in After-markets with Limited Commitment*

In this section we formalise one of the possible rationales for anti-competitive tying in after-markets set out in Section 4.3.6, by relying on an adapted version of Borenstein et al. (2000). We first describe the model in such a way as to reproduce the Chicago School's main result, then use it to show how changes in the main assumptions – mainly, that a firm is *not* able to credibly commit to future prices – will result in supra-competitive prices in the after-market.

4.4.5.1 The model*

Assume that a firm produces a durable good, say a machine, which lasts two periods. In the second period, the machine has depreciated, but a user may buy $q(p)$ units of service from the monopolist, and this will increase the value of usage, where p is the price of service. We assume that producing the good and service has a constant marginal cost that we take equal to 0 without loss of generality.

At any moment in time, and forever after, there are two cohorts of consumers, one which has bought the machine in the previous period, and who therefore has to decide on how much service to buy, and the other who had bought two periods ago and therefore has to buy a new machine. Consumers are uniformly distributed on a line of unit density, and their utility from using the machine will be:

$$U = S - \tau d - P + \delta [S - \tau d - h + f(q(p)) - pq(p)], \quad (4.31)$$

where d is the ideal variety of the good (or, if one prefers, the distance d from the point 0 where the monopolist is located) τ is a disutility parameter, P is the price of the primary product, the machine, S is the per-period reservation value from using the machine, h measures the depreciation of the machine after one period of usage, $f(q)$ is the consumer surplus from q units of service, and δ is the discount factor (equal for consumers and the firm).

We assume that if the service was sold at marginal cost, then consumers would receive a positive surplus from buying at least some small q , and that if both machine and service were priced at cost, the consumer would prefer to keep the machine and maintain it than to buy a new one. Further, we assume that τ is sufficiently large that not all consumers will buy the good (so that the monopolist's demand is affected by its price decisions).

We shall assume that consumers expect future after-market prices to equal current after-market prices, $p_{t+1} = p_t$, and the focus will be on equilibria where prices of both machines and services are constant over time.

4.4.5.2 A Chicago School setting: credible commitment*

As a first step, we solve the model under the (strong) assumption that the monopolist can credibly commit to charge a certain price p of the service in all current and future periods.

In order to derive the demand faced by the monopolist, note that a consumer located at d will buy if the expected net utility from using the good in both periods is positive: $U \geq 0$, which amounts to:

$$d \leq \frac{S - P + \delta [S - h + f(q(p)) - \pi(p)]}{\tau(1 + \delta)} \equiv \widehat{d}, \quad (4.32)$$

\widehat{d} being the indifferent consumer between buying or not (all those located to her left will buy), and $\pi(p)$ is the per-period profit from service, given we assumed that services have zero cost.

At time $t = 0$, when nobody has ever bought the good, the net present value of profits for the monopolist will be:

$$V_0 = \frac{\widehat{d} [P + \delta \pi(p)]}{1 - \delta}, \quad (4.33)$$

which after substituting (4.32) becomes:

$$V_0 = \frac{\{S - P + \delta [S - h + f(q(p)) - \pi(p)]\} [P + \delta \pi(p)]}{\tau(1 + \delta)(1 - \delta)}. \quad (4.34)$$

The monopolist will want to choose p and P in order to maximise its profit V_0 . Taking the FOC with respect to P will give:

$$P = \frac{S + \delta [S - h + f(q(p))] - 2\delta \pi(p)}{2}, \quad (4.35)$$

and from replacing this value back into (4.34) we obtain:

$$V_0 = \frac{\{S + \delta [S - h + f(q(p))]\}^2}{4\tau(1 + \delta)(1 - \delta)}. \quad (4.36)$$

To maximise its profits, the monopolist will therefore need to solve:

$$\frac{\partial V_0}{\partial p} = \frac{2\{S + \delta[S - h + f(q(p))]\} \delta f'(q)q'(p)}{4\tau(1 + \delta)(1 - \delta)} = 0. \quad (4.37)$$

We can therefore immediately see that setting the after-market price equal to its cost is optimal for the monopolist. Indeed, at $p = 0$ (recall we have assumed $c = 0$ without loss of generality), the consumer surplus associated to service, $f(q(p))$, will be maximised, hence $f'(q)q'(p) = 0$.

Therefore, in a world in which the monopolist can credibly commit to set a certain level of service price p and stick to it forever, then it will set it at the competitive level. This is nothing other than the Chicago School's result.

The incentive to renegotiate the after-market price (a 'surprise theory')

To show there exists an incentive to change the price of services, that is, that if the commitment was not credible the monopolist would not stick to the promise of setting p at cost, consider what happens at time $t = 1$ after at $t = 0$ the prices were set as just found – that is, $p = 0$ and P as given by (4.35), which after substitution becomes:

$$P = \frac{S + \delta[S - h + f(q(0))]}{2}. \quad (4.38)$$

where we have used $\pi(p = 0) = 0$ (as there are no after-market profits when selling at a price equal to cost).

In turn, by replacing into (4.32) we obtain the demand at period $t = 0$, as:

$$\widehat{d}_0 = \frac{S + \delta[S - h + f(q(0))]}{2\tau(1 + \delta)}, \quad (4.39)$$

Therefore, at $t = 1$, the monopolist will have the following expected stream of profits:

$$V_1 = \widehat{d}_0\pi(p) + \frac{\widehat{d}[P + \delta\pi(p)]}{1 - \delta}, \quad (4.40)$$

where the second term on the right-hand side of expression (4.40) is nothing other than the original expression for V_0 . The maximisation of V_1 with respect to the machine price P will entail setting $\partial V_1/\partial P = 0$. However, since P does not appear in the first term, this will amount to setting $\partial V_0/\partial P = 0$. Hence, P will be given by expression (4.35), and after substituting the second term of (4.40), corresponding to V_0 , will be

given by expression (4.36), so that:

$$V_1 = \frac{S + \delta [S - h + f(q(0))]}{2\tau(1 + \delta)} \pi(p) + \frac{\{S + \delta [S - h + f(q(p))]\}^2}{4\tau(1 + \delta)(1 - \delta)}. \quad (4.41)$$

Maximising V_1 will entail solving:

$$\begin{aligned} \frac{\partial V_1}{\partial p} &= \frac{S + \delta [S - h + f(q(0))]}{2\tau(1 + \delta)} \pi'(p) \\ &\quad + \frac{2 \{S + \delta [S - h + f(q(p))]\} \delta f'(q) q'(p)}{4\tau(1 + \delta)(1 - \delta)} \\ &= 0. \end{aligned} \quad (4.42)$$

It is now straightforward that $p = 0$ cannot be an equilibrium. Indeed, by setting the service price at cost, the second term will disappear: as we observed above, when $p = 0$ the service surplus is maximised, so $f'(q)q'(p) = 0$. Hence, evaluated at $p = 0$, we have $\partial V_1 / \partial p = \widehat{d}_0 \pi'(p = 0) > 0$, since it is obvious that service profits will increase by setting a higher p .

This shows that – after having communicated that it would set $p = 0$ forever, and if it could break that promise, the seller would charge a higher after-market price. In other words, it would want to ‘surprise’ the first generation of buyers by setting a higher after-market price. Equivalently, if it had promised to keep an open system, it would have an incentive to exclude other sellers of services.

4.4.5.3 Limited commitment theory*

Borenstein et al. (2000) show that a firm has an incentive to set $p > c$ in a model where there is no ‘first generation’ of consumers. They use the same overlapping generation model as above and look at what happens at any point of time t . At t , there will be an installed base d_{ib} (we do not know what expectations these consumers had) and the seller’s profits will be:

$$V_t = d_{ib} \pi(p) + \frac{\widehat{d} [P + \delta \pi(p)]}{1 - \delta}, \quad (4.43)$$

where the last term is the usual expression for V_0 , and the maximisation problem for the monopolist is the same as when it had profits

V_t , so:

$$\frac{\partial V_t}{\partial p} = d_{ib}\pi'(p) + \frac{2\{S + \delta[S - h + f(q(p))]\} \delta f'(q)q'(p)}{4\tau(1 + \delta)(1 - \delta)} = 0, \quad (4.44)$$

which can also be rewritten as:

$$\frac{\partial V_t}{\partial p} = d_{ib}\pi'(p) + \widehat{d}(p) \frac{\delta f'(q)q'(p)}{(1 - \delta)} = 0, \quad (4.45)$$

after recalling the expression for $\widehat{d}(p)$ given by (4.39).

At the steady state, it must be $d_{ib} = \widehat{d}(p)$, hence:

$$\frac{\partial V_t}{\partial p} = \widehat{d}(p) \left[\pi'(p) + \frac{\delta f'(q)q'(p)}{(1 - \delta)} \right] = 0. \quad (4.46)$$

As we know, at $p = 0$ we have $f'(q)q'(p) = 0$. So it becomes clear that setting a price p for the service equal to cost (which is without loss of generality equal to 0) cannot be optimal. Indeed:

$$\frac{\partial V_t}{\partial p}(p = 0) = \widehat{d}(0)\pi'(0) > 0. \quad (4.47)$$

In other words, the firm will set the same aftermarket price p above cost in all periods, and all consumers will correctly anticipate this in their decisions.

4.4.5.4 Competition in the primary market*

The main purpose of Borenstein et al. (2000) is to show that the incentive to set after-market prices above costs is still there when firms compete in the primary market. To this extent, they analyse the (Hotelling) duopoly version of the model above, assuming that the two firms produce differentiated primary goods, but then each of them is a monopolist in the services of its own machine. The mathematics of the duopoly model are slightly different, but the logic is the same as above. If it could credibly commit to future after-market prices, a firm would have an incentive to set $p = c$ because, by doing so, it would raise the total expected consumer surplus for its machine and could increase its machine price and/or the demand for its machines (other things equal, consumers would prefer a system with lower service prices). But whenever facing an installed base, a duopolist will want to raise p above cost: like in the model just analysed, at $p = c$ an increase in p would increase service profits by a first-order effect, whereas it would decrease profits from current and future consumers by a second-order effect.

4.5 From Theory to Practice

4.5.1 Some General Principles

As we have noted throughout the chapter, tying and bundling (including mixed bundling, that is, bundled rebates) are common business practices used by dominant and non-dominant firms alike. As such, as a first step in any investigation, it would be helpful to understand any pro-competitive or efficiency rationales for them (see Section 4.2.1) or objective justification (for example in relation to health and safety considerations or protection of intellectual property). In some cases, such an assessment may already enable a competition authority to reach a view that a case should not even be opened, for example if the tying consists of a product innovation that is likely to increase consumer welfare.

As part of an initial assessment, it may also be helpful to understand whether a firm's conduct might be driven by a price discrimination motive that seeks to address customer groups with different preferences for different products (or their combinations) – see Section 4.2.2.

As discussed in this chapter, under certain conditions, tying, bundling and bundled rebates may raise anti-competitive effects. The theories of harm that may underpin these practices might be of a different nature. Further below, we will therefore first consider those related to tying (including in the context of after-markets) and (pure) bundling; and, separately, to bundled rebates.

Regardless of the precise theory of harm that one may investigate, an important element for the assessment is whether the alleged infringer is dominant (that is, it has significant market power) in the tying market.⁶⁸ This is a legal requirement in most jurisdictions (as far as unilateral conduct provisions are concerned)⁶⁹ but it is also consistent with a number of the

⁶⁸ Cases related to after-markets would deserve some qualification, as we discuss further below in this section.

⁶⁹ We are aware that both in the US and in the EU (and quite possibly beyond) tying may also be caught by competition law as a potential anti-competitive agreement between firms as opposed to anti-competitive unilateral conduct. Indeed, a number of tying (and bundling) cases in the US were assessed under Section 1 of the Sherman Act (see our discussion of *Loew's* in Section 4.6, for example; *Microsoft* and *PeaceHealth*, just to name two further US cases, also had some Sherman Act Section 1 aspects to them). In the EU, the European Commission's Guidelines on Vertical Restraints (2010) note that tying may constitute a vertical restraint falling under Article 101 TFEU where it results in a single-branding type of obligation. These guidelines indicate that the Commission is unlikely to challenge instances of tying where 'the market share of the supplier, on both the market of the tied product and the market of the tying product, and the market share of the buyer, on the relevant upstream markets, do not exceed 30%' (see para. 218).

economic models we have reviewed in this chapter (such as Whinston, 1990, Carlton and Waldman, 2002, Hurkens et al., 2016). We believe that it is highly unlikely that tying by a firm without significant market power can lead to material anti-competitive effects.^{70,71}

4.5.2 Tying and Bundling Theories of Harm

As with our discussion of policy implications in other chapters, a finding of anti-competitive conduct (in this case, through tying) must be underpinned by a clear and robust theory of harm. In other words, one needs to understand the rationale behind the tying, and whether this rationale should be expected – based on clear mechanisms, such as those present in some key economic models – to ultimately lead to actual or likely foreclosure effects which are material.

In this chapter, we have considered four broad categories of theories of harm.

First, there exist theories of harm based on so-called *imperfect rent extraction*. Put simply, if the alleged infringer faced a more efficient rival in the supply of the tied product (or where the rival produces a different or a higher quality version of the tied good, either currently or in the future, or even a new complementary product in the future), it would typically want to allow it to serve customers and then to extract the rival's

⁷⁰ Tying may also be used as a defensive strategy, as an old UK case, *Ilford*, shows. (See Monopolies Commission, *Colour Film – A Report on the Supply and Processing of Colour Film*, H.M.S.O. No. 1 (1966).) There, a non-dominant firm, Ilford, was forced to discontinue its tying practices in 1966, following complaints by smaller competitors in the processing business. Ilford used to tie the sale of colour film together with its processing. It had about 15 per cent of the UK market for film sales (Kodak was the market leader with an 80 per cent market share). Ilford had found the tying business model an effective way to counter the indirect network effects brought about by Kodak's conduct: independent processors mostly used processing technology compatible with the market leader, but which was incompatible with Ilford's films. After losing its case, Ilford abandoned its business practice and was then forced out of the market by 1968. See Motta (2004: section 7.3.2.4) for further discussion.

⁷¹ From a legal perspective, one also needs to show that the products under scrutiny are distinct (and not part of a unique, aggregate product). This is required so that tying or bundling are meaningful conducts to assess; otherwise, if what were believed to be different products constituted instead one single product, it may be more appropriate to investigate a different type of conduct, for example predatory pricing concerning the entire bundle. In a sense in the different models we reviewed, by referring to product *A* and to product *B*, and to different strategies that a firm may adopt in relation to these products, we were implicitly positing that these products were distinct. In practice, determining whether products are distinct or not may not always be straightforward and a number of cases involved an extensive debate on this matter (see Sections 4.6 and 4.7).

efficiency rent (or the other rents available more generally) through the price of the tying product; this would typically be more profitable than foreclosing the rival. If, however, the alleged infringer is prevented from fully extracting such efficiency rent, an incentive to foreclose may arise, and tying could potentially be a tool to achieve that goal. In this case, it is therefore crucial to be able to explain why the alleged infringer is unable to extract the rival's efficiency rent (as well as ensuring that the rival is at least as efficient as the alleged infringer or is indeed able to produce superior versions of the products). For example, there might be constraints on the price of tying product (due to regulation, say); or on the ability to use price structures that are more complex than simple unit prices in a context where the tying product and the tied product are independent (as opposed to complementary); or on the ability to write contracts that are contingent on the development of future upgrades or products.

Second, there are some theories of harm where tying acts as a device to *commit to aggressive conduct as a way to discourage entry* (originally due to a model by Whinston, 1990). A key requirement for this theory of harm is to show that the alleged infringer's commitment to tie (or bundle) is credible; that is, the tying decision cannot be reversed. A related, critical point of this literature is the following: if bundling is profitable only to the extent that it will result in entry-deterrence (this is why a credible commitment to bundle is necessary), then the dominant firm should engage in it only when it has a fairly accurate knowledge about the cost structure of the rival and the likely profits it would obtain if it engaged in bundling. In other words, this theory seems to have a pretty strong informational requirement. Recent economic research shows that in certain circumstances bundling may result in more aggressive competition and may be a profitable strategy even if entry did take place. (See Hurkens et al., 2016 and Peitz, 2008.) However, the same research shows that in slightly different settings, the effect of tying may be to soften competition, rather than to make it fiercer. We have also seen how the welfare implications from tying in some of these models (for example in Sections 4.3.3.2 and 4.3.3.3) may be ambiguous. Absent at this stage a clear understanding of which variables could indicate the likely direction of the strategic effect of bundling, it might be unwise to build an exclusionary case on this 'strategic' theory of harm.⁷²

⁷² An exception could be a situation where internal documents indicated a coherent strategy aimed at using bundling with the sole purpose of excluding rivals, but this is likely to be a very rare occurrence.

Third, we have considered a fairly specific theory of harm (originally due to a model by Choi and Stefanadis, 2001) where tying (pure bundling, more precisely) acts as a way to *raise entry barriers in a context where entrants need to invest (for example, in R&D) in every market* where the alleged infringer is active and only have some probability of that investment being successful and thus putting them in a position to challenge the incumbent. By tying different products together, the incumbent manages to make the degree of competition in one market dependent not only on the success of the entrant's investment with respect to that market, but also on the success of other entrants in other markets. In practice, this makes it more difficult (or more precisely, less probable) for competition to develop in a given market than in the absence of tying. Under this theory of harm, tying needs to be irreversible. Further, the incumbent needs to be able to commit to it before the entrants make their investment decisions. In sum, this theory of harm too relies on fairly specific (potentially unlikely) conditions and it would probably be difficult to be able to demonstrate that tying was undertaken with a strategic motive;⁷³ as such, we would advise particular caution with it as well.

Fourth, we have presented a theory of harm (due to Carlton and Waldman, 2002) which is based on the goal to *protect monopoly power in the tying market*. The key requirements for such a theory of harm to be robust include the presence of significant scale or scope economies, either on the supply-side (for example, in the form of large fixed costs of entry) or on the demand-side (in the form of network effects or network externalities). Crucially, for one to be able to invoke this theory of harm, one has to be able to show a clear linkage between the ability to enter into the tied market currently and that to enter the tying market at a future stage. In other words, one needs to be able to show that the rival needs sufficient scale in relation to the tied product currently so as to be able to enter the tying product market in the future. A further requirement, in the case of the version of this theory of harm that relies on scale economies on the supply-side, is that the alleged infringer needs to be able to commit to the tying of the two products: that is, it has to be able to be credible in its policy that – should there be entry in the tied product market – it will not sell the tying product on a stand-alone basis.

⁷³ Again, an exception may potentially be if internal documents indicated a coherent strategy aimed at using bundling with the sole purpose of excluding rivals.

Some remarks Before we turn to issues specific to tying in after-markets, we conclude with three remarks that are likely to be relevant across the various theories of harm just considered.

First, as discussed above, a typical requirement for tying to raise competition concerns is that *a firm needs to be able to commit* to it. In this regard, technical tying (for example, where different products are physically assembled together) is more likely to raise concerns than contractual tying, which can typically be reversed relatively swiftly.⁷⁴ Note, however, that technical tying is often associated with product innovations (see Section 4.2.1) and therefore with possible pro-competitive effects. By contrast, where tying is purely contractual (and there is no credible commitment not to sell the products on a stand-alone basis), several theories of harm become inapplicable, so that regardless of the existence of efficiencies or objective justifications, there are few robust ways in which tying conduct may raise significant competition concerns.

Second, *product differentiation* in the tied market makes tying less likely to have a foreclosure effect, all else equal. This is because some consumers may decide not to buy the bundle when the rival's variant of the tied product is not available, which limits the profits that the incumbent earns when it engages in bundling. This is a conclusion that can be drawn from a number of the models we reviewed in this chapter, including Whinston (1990).

Third, an important driver of any ability of tying to raise likely significant foreclosure effects in the tied market is the *coverage of the tied market* (in other words, the tied percentage of total sales on the market of the tied product). Even if the incumbent is dominant in the tying market, it may still be the case that only a small proportion of the tied market is affected by such tying. For example, a bank that is dominant in the supply of personal current (or checking) accounts⁷⁵ may offer car loans only to its personal current account customers. To the extent that the proportion of car loan customers who buy this product from such bank is low relative to the overall number of car loans (for example because there are many successful stand-alone providers of this product), it is difficult to see how the tying conduct could foreclose competition in the tied market. One exception to this argument may arise in the context of

⁷⁴ The European Commission also makes this point in its Guidance Paper on its enforcement priorities in relation to exclusionary practices. See European Commission (2009).

⁷⁵ We are assuming for the sake of this example that this could constitute a relevant market for antitrust purposes.

some business-to-business relationships where the low proportion of tied buyers amount to ‘key buyers’. For example, these may be buyers with a particularly strong growth potential, or which are crucial to sellers in some other respect (for example because they have a distribution network that particularly suits a rival or a new entrant). In this case the potential for anti-competitive concerns to arise is greater than otherwise (see also our discussion in Chapter 2 of the role of key buyers in the context of the assessment of single-product rebates).

Tying with after-markets (requirement tying) We provided a discussion of some of the key antitrust issues in relation to after-markets in Section 4.3.6. As discussed there, to the extent that there is effective competition in the primary market, it is difficult to see how conduct in an after-market (such as requirement tying by a branded good manufacturer of a primary product together with after-sales) may raise competition concerns. In fact, if competition in the primary market is effective, it would probably be more appropriate to define a system market that encompasses both the primary product and after-markets; this would rule out any dominance of a given firm at the level of a specific after-market (for example the spare parts of a given brand of a primary product) and thus tying as an abusive conduct would no longer be a meaningful concept.

With reference to the academic literature, we have also seen how there may be some limited exceptions to the above argument. However, any theory of harm which may explain why a firm may want to exclude rivals from its after-markets would depend on some specific conditions applying (see Section 4.3.6) and any consumer harm is likely in any event to be limited. In fact, we have noted that antitrust action may even be counter-productive, for example if intervention pushes after-market prices downwards where firms have already competed away after-market rents through low prices of the primary product.

4.5.3 Bundled Rebates Theory of Harm

As noted in Section 4.3.5, the only robust theory of harm in relation to bundled rebates is, to our knowledge, one that is essentially equivalent to a predatory one with respect to the ‘tied’ product (which we have typically called product *B* throughout this chapter, and where the incumbent might not be dominant). Put otherwise, differently from tying, we are not aware of robust theories of harm where the interplay between different markets may underpin the anti-competitive nature of bundled rebates.

As such, our view is that the assessment of bundled rebates should broadly follow that of predatory allegations (see our discussion in Chapter 1 for details). Therefore, for bundled rebates to be potentially anti-competitive, we would expect the incumbent to at least have some market power with regard to product *B*. We would also expect there to be significant barriers to entry and/or to expansion, otherwise the incumbent would be unlikely to recoup any profit sacrifice in future periods, and this would beg the question as to why the bundled rebates were offered in the first place (absent any efficiencies, of course).

Price-cost test If such a predatory-like theory of harm appears to be robust based on the facts of a case, it would then be important to perform a price-cost test. To perform this test, one would need to consider the prices of the bundled offer and of the stand-alone products, and then allocate the whole discount associated with the bundled offer to the product where the incumbent faces competition (product *B*, in the jargon of this chapter).⁷⁶ The resulting ‘effective’ price of product *B* should then be compared to a reasonable cost benchmark. Given the similarity with predation cases, for a discussion on how to apply this test, we refer the reader to the discussion in Chapter 1. In particular, it may be appropriate to consider two benchmarks for the cost of product *B*. The upper bound may consist of average total costs (‘ATC’), or long-run average incremental costs (‘LRAIC’). The lower bound may consist of average variable costs (‘AVC’), average avoidable costs (‘AAC’) or short-run average incremental costs (‘SRAIC’). Then, if the effective price of product *B* is higher than the upper bound of such cost range, the allegation of anti-competitive conduct should be dismissed. If it is lower than the lower bound of the cost range, then this may constitute *prima facie* evidence of anti-competitive conduct. Finally, if the effective price falls between the bounds of the cost benchmarks, a deeper assessment of other market conditions and conduct may be required, with a focus on determining the existence of any precise mechanism that would explain an incentive of the incumbent to

⁷⁶ For instance, if products *A* and *B* bought separately are priced respectively at \$10 and \$5, while a bundle with both products is priced at \$13, then the effective price of product *B* would be \$3 (= \$13 – \$10). In other situations, for instance when they do not have a one-to-one relationship, the two products may be priced separately in the bundle. In this case, one can compute the effective price of product *B* by computing the discount lost on product *A* when the bundle is not purchased (as the difference between the individual price of product *A* and its price in the bundle) and by attributing that discount to product *B*. Similar principles were proposed in relation to the ‘discount attribution standard’ in *PeaceHealth*, which we review in Section 4.6.

exclude rivals in the supply of product *B* through some profit sacrifice (see Chapter 1).

Some remarks In any event, we would caution against a mechanical use of the above-mentioned price-cost test and we would make the following remarks.

First, consistent with all other policy implications across this book, we stress the robustness of the theory of harm as the key factor in the assessment of unilateral conduct.

Second, as set out in Section 4.2.2, there may be instances in which a firm – as a result of mixed bundling or bundled rebates – sets an effective price for an individual product below cost without necessarily an exclusionary motive, but rather to address customer groups with different valuations for different products more effectively and we know that such a ‘price discrimination’ rationale has ambiguous welfare consequences.

Third, similar to our discussion in relation to tying, the extent of the coverage of the bundled rebates is likely to be an important driver of any foreclosure potential of that pricing strategy. That is, all else equal, the smaller the proportion of the ‘tied’ market that is affected by the bundled rebates, the weaker the potentially exclusionary effect of the bundled rebate. Take the example of a firm which is dominant in grocery retailing and which has no market power in the retailing of car fuel. Such a firm may provide car fuel discounts to customers who are also buying their groceries from it. However, it is difficult to conjecture that these bundled rebates have the objective or the ability to foreclose established competitors in car fuel retailing, potentially even if they lead to prices below some cost benchmark.⁷⁷ In Section 4.7, we will see in our discussion of *IDEXX* how the low coverage of some bundled discounts was one of the key factors that led the UK competition authority to dismiss allegations of anti-competitive bundled rebates, even in the presence of some evidence of below-cost pricing. One potential exception to this argument is, like in the context of tying, where the buyers affected by the bundled rebates are ‘key buyers’, for example with a particularly strong growth potential, or which are crucial to sellers in some other respect; and if supported by a coherent theory of harm which explains in which sense these buyers are so crucial, then there may be reason for anti-competitive concerns.

⁷⁷ Another example may be a dominant bank offering a discount on home insurance products only to its banking customers, where the bank has no market power in relation to home insurance (or non-life insurance more generally) and faces effective competition for the provision of this product.

4.6 Case-law

In both jurisdictions, competition authorities and courts have typically assessed tying and bundled rebates under separate lenses. Therefore, in each of the following two sections (covering, respectively, the US and the EU), we begin with a review of the key tying cases, followed by an overview of bundled rebates ones.

4.6.1 United States

Tying Most of the early US tying cases concerned the allegedly anti-competitive use of patents through patent tie-ins: firms with a patent over a certain machine would force customers to purchase any related materials from that same company, arguing that any other usage of the machine would infringe the patent.⁷⁸ As a response, the courts took a very pro-active stance against this behaviour: tying practices led to ‘automatic condemnation’ (see Areeda and Hovenkamp (2003: para. 1701c), where the authors cite a large number of cases). The theory of harm implicitly put forward in most of these judgments was that of ‘leveraging market power’ from the tying product market to the tied product market. While these early patent tie-in cases largely entailed primary market – after-market relationships (discussed in Section 4.3.6), these cases were not interpreted in this spirit; this would take a few decades, as we shall see below. Indeed, in the early judgments, there was no assessment of competition in the ‘primary market’ and the defendant was simply assumed to have monopoly power in the supply of the tying product on the basis of the patent it held on this product.

In *Motion Picture Patents*, the holder of a patent over film projectors (MPPC) complained that its customers failed to comply with its selling condition that these projectors should only be used with films produced by MPPC’s licencees (threatening the projector’s confiscation otherwise). The Supreme Court, in 1917, dismissed the action brought forward by the incumbent firm.

In *IBM*, in 1936, the incumbent (IBM) was found to have illegally leased its tabulating machines upon the condition that customers obtained the punch cards from IBM itself. As Carlton and Waldman (2002), as well as Areeda and Hovenkamp (2003: para. 1720) noted, the most plausible explanation for IBM’s behaviour was the desire to price discriminate; tying

⁷⁸ See, for example, Areeda and Hovenkamp (2003).

would therefore be deployed as a metering device, in the fashion we discussed in Section 4.2.2.

Eleven years later, in *International Salt*, the Supreme Court ruled against International Salt for tying its salt to the lease of two of its patented machines, Lixator and Saltomat, which processed the salt further for other industrial uses. International Salt was found to have market power in the tying markets (the two machines) purely on the basis that they were patented. The defendant claimed an efficiency defence, justifying its tie with the necessity to guarantee high quality standards for the salt used in the machines, to avoid technical damage. The Court rejected this justification.

In *Northern Pacific Railway* the eponymous railway company sold and leased land next to its rail lines upon the condition that the purchasers (or lessees) used Northern Pacific's rail lines to ship goods produced over that land (most of which was grazing land). The Supreme Court found that this tying agreement constituted a restraint of trade. The tying product (over which Northern Pacific was found to have 'substantial economic power') was land. Cummings and Ruhter (1979) discard the leverage motive of tying on the basis that Northern Pacific simply did not have market power over land: even in the State where it was strongest against its competitors, Montana, it owned less than 2 per cent of the total land (and about 2.5 per cent of timber land). Instead, an interesting feature of the case was that the railroad company forced the tie only when its rail freight rates were equal to or lower than those of its competitors. An explanation of the tie, given by Cummings and Ruhter (1979), is based on tacit collusion: this clause could help Northern Pacific monitor its railroad competitors' rates. The authors suggest that the evidence provided during the trial supported this interpretation.⁷⁹

In *Loew's*, in 1962, the Supreme Court found that the practice by distributors of copyrighted feature motion picture films for television of licensing such films to television broadcasting stations only in bundles (also containing licences to films the stations did not want) constituted

⁷⁹ The authors write: 'Virtually all the depositions reveal that Northern Pacific officials regularly visited land users to discuss shipments. When a buyer or lessee did not use Northern Pacific, the contract compelled him to disclose the lower rates or better service available elsewhere. If he was uncooperative *in reporting*, the [contract] clauses would have given Northern Pacific grounds for legal action, for refusing to renew a lease, or for not accepting an application for a new purchase or lease. In addition, some timber contracts required these buyers to keep records of production and shipments and to permit Northern Pacific agents regular access to those records.'

unlawful tying in violation of Section 1 of the Sherman Act (that is, this case was not assessed under the monopolisation provisions). The Court took issue with what it called ‘block bookings’, which we have earlier in this chapter referred to as *full-line forcing*. The Court also imposed that the decrees entered to in this case by a lower court with the defendants should be amended as to oblige the distributors to offer the licences on an individual basis; further, if distributors additionally offered bundles, they would be prevented from charging a given licence differently depending on whether it was bought individually or as part of a bundle (unless this could be justified on the basis of legitimate cost considerations).

In many of these previous cases, among others, a so-called ‘goodwill’ efficiency defence was put forward: the defendants claimed that the tie was justified to protect their brand image and to guarantee consistently high quality. Such defence had always been rejected by the courts. Thus, when *Jerrold* was handed down, in 1960, it constituted an exception to the earlier strict approach. *Jerrold* was in the television antennae business, which was at its infancy in the late 1950s. To protect the well-functioning of its products, it forced its customers to accept only its own installation and maintenance services and banned any equipment addition from any other company. The District Court accepted the goodwill defence on the basis that the product was very new and that the firm would have run a high risk of damaging its reputation at an early stage, if there was any malfunctioning. The continuation of the clause over time, though, was deemed unlawful.

Another landmark tying case was *Jefferson Parish*. This related to Jefferson Parish Hospital’s (‘JPH’s’) agreement with an anaesthesiologists group, Roux and Associates (‘RA’).⁸⁰ Under this agreement, any (of the four) anaesthesiologists belonging to RA would only practice at JPH; and no anaesthesiologist could practice at JPH unless he or she belonged to RA. As such, JPH was found to tie the provision of hospital services to patients to that of anaesthesiologist services: JPH patients were prevented from sourcing anaesthesiologist services elsewhere. The Appeals Court judgment had found liability on the basis of a *per se* approach. One of the necessary conditions for the *per se* approach to be applicable was the existence of ‘sufficient economic power’ in the tying market to coerce a

⁸⁰ Here we draw from Nalebuff and Majerus (2003), who summarise and provide an economic assessment of the case that corroborates the conclusions drawn by the Supreme Court.

purchase in the tied product market.⁸¹ In this specific case, the Supreme Court reasoned that the hospital did not have sufficient market power in the relevant geographic market once correctly defined (it had about 30 per cent of the patients). Put otherwise, patients may have opted for a different hospital (with different anaesthesiologists) if they believed that the price or the quality of hospital or anaesthesiologist services at JPH were not competitive. Further, the Supreme Court found that the agreement between JPH and RA had not led to any increase in price or a decrease in the quality of anaesthesiological services. As such, the Supreme Court overturned the Appeals Court's judgment.

Yet a broad prohibition of tying practices was still effective, as *Kodak* showed. This case concerned copying machines, spare parts and the provision of maintenance services.⁸² Kodak's competitors had entered an after-market (maintenance services, in this case) by purchasing spare parts from Kodak and other manufacturers, and then providing their own maintenance services to equipment owners (including those of Kodak's products). Kodak responded by tying its spare parts to its own maintenance services. The Supreme Court found this behaviour anti-competitive, as in its view Kodak possessed 'appreciable economic power in the tying market', which it defined as the supply of spare parts used for Kodak's copiers. On the other hand, the Court agreed with Kodak that Kodak did not have market power in the primary market for copiers. Research by Nalebuff and Majerus (2003) show that, in the period from 1986 to 1991, Kodak had a market share of 5 per cent in the case of mid-volume copiers and of 20–35 per cent in the case of high-end copiers (whereas Xerox supplied just under 40 per cent of mid-volume copiers and just over 50 per cent of high-end ones). Kodak defended itself by claiming that if it raised the price of spare parts above the competitive level, a significant proportion of consumers would migrate to an alternative equipment provider (together with alternative, possibly cheaper, provision of spare parts and servicing). The Court rejected this defence for the following reasons. First, this claim fell into the cellophane fallacy (a firm with market power would have already raised its price up to the profit-maximising level, so that further price increases would have been unprofitable). Second, the Supreme Court rejected Kodak's (life-cycle pricing) argument that

⁸¹ The other three were: (i) there are two separate products or services involved; (ii) the sale of one product or service is conditioned upon the purchase of the other; and (iii) not an insubstantial amount of interstate trade in the tied product is affected.

⁸² See Borenstein et al. (1995), Goldfine and Vorrasi (2004), Klein (1996), Nalebuff and Majerus (2003) and Shapiro and Teece (1994) for further discussion of this case.

customers would assess the value of the overall package of copier, spare parts and maintenance services over time. The Court pointed to the potential irrationality and ignorance of new customers (who would not take the price of maintenance services into account when purchasing a copier). Third, the Court emphasised the high switching costs and lock-in effects for existing customers. This was a split decision. The three dissenting justices made two very sound arguments, from an economic perspective: first, that it was quite odd to keep applying *per se* illegality principles, typically adopted for unequivocally harmful practices like price-fixing, in areas like tying where there is some consensus among economists that this may also entail benefits. Second, the majority view of the Supreme Court discounted the state of competition in the primary market (copiers) and purely focused on intrabrand competition (that is, price competition over Kodak products), neglecting interbrand competition which appeared to be healthy.

Yet the tide began changing as Goldfine and Vorrasi (2004) point out, at least in the case of the tying of primary and after-market products. Following a review of many lower court judgments handed down after *Kodak*, the authors conclude that the courts, in practice, interpreted *Kodak* as follows: firms would be found to have appreciable economic power in the supply of the aftersale product only if three cumulative conditions were met. First, consumers faced high switching costs. Second, consumers lacked sufficient information on the whole life cost of the product before purchasing the primary product. And third, once consumers were 'locked-in', the firm engaged in opportunistic conduct to 'exploit' those consumers. Given these conditions, the authors noted, it became much more difficult for plaintiffs to prove that firms tying primary and after-market products were acting illegally.

It can be argued that the relaxation of antitrust enforcement in the area of tying affected more than just cases involving after-markets. Following the approach set out above in the context of *Jefferson Parish*, defendants had the possibility of pointing towards the existence of market conditions that were inconsistent with those required for a Court to apply a *per se* approach to the assessment of tying (which would imply the tying would be illegal). If the defendants' arguments were successful, the tying practice would have had to be assessed under a *rule of reason*, which would make a tying practice much less likely to be found illegal.

Some cases inevitably depended on the precise boundaries of the relevant market(s). In *Queen City Pizza*, Domino's Pizza was taken to court by 11 of its franchisees for tying the franchise to the purchase of ingredients

from Domino. The District Court dismissed the claim on the basis that the plaintiffs defined the market erroneously ('ingredients, supplies, materials, and distribution services used by and in the operation of Domino's pizza stores'). The Appeals Court agreed. From an economic perspective, there was a limit to which Domino's Pizza could exercise its market power (if any): regardless of any tying or exclusive dealing, should Domino's Pizza's final retail price be too high, consumers would likely switch to other restaurants.

Microsoft, decided back in 2001, represented an important precedent for US antitrust enforcement on tying, giving prominence to innovation aspects and a need for a rule of reason assessment. This case involved the technological tying of the Internet Explorer web browser into Microsoft's Windows operating system. Microsoft was found to have 95 per cent of the market for Intel-compatible PC operating systems; this, combined with the barriers to entry found in the applications market, led the District Court to conclude that Microsoft had monopoly power and found the tying *per se* unlawful. But the Appeals Court rejected such approach and remanded the matter to the District Court for a rule of reason assessment. This was because – the Appeals Court argued – there could be efficiencies, with consumers valuing an integrated product more than the sum of its parts; and these pro-competitive effects could in principle outweigh any anti-competitive effects arising from tying. Microsoft eventually settled the case with the US Department of Justice. We discuss this case in more detail in Section 4.7.

The process started with *Kodak* saw a great leap forward with the Supreme Court's judgment in *Independent Ink*. As we noted above, the jurisprudence on tying was developing by narrowing down the set of conditions that would justify a *per se* prohibition of tying. In *Independent Ink*, the Supreme Court clearly stated that, in a tying case involving intellectual property, plaintiffs need to show that the patent holder possesses substantial market power in the tying market. It also clarified that holding a patent *per se* is not sufficient for such a finding. The case arose out of patent infringement litigation in the context of a specialised technology used to print bar codes on cartons. Illinois Tool Works tied its (patented) printers to its own (non-patented) ink. Illinois Tool Works filed a patent infringement suit when it learned that Independent Ink started marketing its own ink for use on Illinois Works Tool's printers. Independent Ink responded by filing an antitrust lawsuit, on the basis of unlawful tying. The District Court dismissed the claim, arguing that sound economics would require a deeper investigation of the defendant's market power in the tying

market, and not simply a market power presumption based on it holding a patent over its own printer. The Appeals Court reversed, somewhat reluctantly, stating that the long case-law supporting the presumption of market power by a patent holder could only be overturned by the Supreme Court. This is precisely what the Supreme Court did, reversing the Appeals Court's decision and clearly stating:

Congress, the antitrust enforcement agencies, and most economists have all reached the conclusion that a patent does not necessarily confer market power upon the patentee. Today, we reach the same conclusion, and therefore hold that, in all cases involving a tying arrangement, the plaintiff must prove that the defendant has market power in the tying product. (Paras 45–6)

Bundled rebates US jurisprudence of bundled rebates followed a very different trajectory from that of tying. This was probably due to the feasibility of price-cost tests, which the courts appeared to prefer applying, in line with jurisprudence on predation and single-product rebates. (By contrast, such tests are generally not meaningful in the case of pure bundling.) Perhaps as a result of this, as well as due to the traditional observation that lower prices generally benefit consumers, there has never been a *per se* prohibition of bundled rebates in the US. Nevertheless, *LePage's*, which we discuss below, proved controversial, as it departed from the more traditional US stance on discounts.

We next briefly review what we understand to be the four most prominent cases. The first to reach the courts was *SmithKline*.⁸³ SmithKline and Eli Lilly were selling cephalosporin antibiotics to hospitals, which were deemed to constitute a relevant market. In the early 1970s, Eli Lilly, the dominant firm according to the courts, began offering rebates on its four patented antibiotics within this market, as well as on Kefzol, its *off-patent* cefazolin drug (that is, Eli Lilly faced competition over this specific antibiotic). In 1975, Eli Lilly modified the existing rebate structure: it reduced the level of rebates, but at the same time it introduced a new 3 per cent rebate that was applicable across all cephalosporin antibiotics purchased from Eli Lilly, provided customers (hospitals) met individual volume targets on any three out of the five antibiotics. Two of those five medicines – both patented – had very low demand by 1975, according to the records; the other two patented drugs jointly covered about 68 to 86 per cent of hospital needs of cephalosporin antibiotics between 1973 and 1975; in the off-patent segment, cefazolin and its substitutes (including

⁸³ Here we also draw from Greenlee et al. (2008) and Kobayashi (2005).

SmithKline's Ancef) received significant demand. Based on such demand pattern, it appeared that meeting the target on cefazolin (the only segment subject to competition at the time) was a *de facto* necessary condition to qualify for the multi-product rebate. The courts pointed to evidence that hospitals were choosing Ancef over other cephalosporins offered by Eli Lilly: for instance, Keflin (one of the two patented cephalosporins with large hospital demand) was replaced by Ancef over 80 per cent of the time. Parties agreed that Eli Lilly's costs for its cefazolin were lower than SmithKline's, but the latter believed that it could bring its costs down to Eli Lilly's level, as soon as it reached sufficient sales volumes.

The District Court relied on a hypothetical as-efficient competitor test: it received evidence that if SmithKline had Eli Lilly's cost structure and it offered rebates on Ancef to make it competitive with Eli Lilly's bundled rebate, SmithKline would have earned a return of 4 per cent on average hospital accounts (and -2.7 per cent in the case of large hospital accounts). These returns, the Court argued, would have made detailing (that is, the promotion of the drug to doctors) not worthwhile, in turn reducing the demand for Ancef. The Appeals Court confirmed the lower court's finding that Eli Lilly's practice amounted to illegal maintenance of monopoly power: in its view, the bundled rebate scheme had an anti-competitive, exclusionary objective targeted at Ancef.

A similar question arose in *Ortho Diagnostic*. According to the information available in the judgments, the facts of the case were broadly as follows. Abbott Labs was in the business of selling, *inter alia*, five complementary tests for the screening of viruses in the blood supply. Ortho only competed on one of these tests (for Hepatitis C). Abbott had a market share ranging from 70 to 90 per cent in the other four markets. Abbott Labs provided strong incentives to the Council of Community Blood Centers in order to buy its products, by offering a bundled rebate conditional on four or five tests being purchased.

The District Court set out the following test to determine the likelihood that a multi-product rebate scheme may be anti-competitive:

Under [the monopolisation provisions] a plaintiff in a case in which a monopolist (1) faces competition on only part of a complementary group of products, (2) offers the products both as a package and individually, and (3) effectively forces its competitors to absorb the differential between the bundled and unbundled prices of the product in which the monopolist has market power, must allege and prove that either (a) the monopolist has priced below its average variable cost or (b) the plaintiff is at least as efficient a producer of the competitive product as the

defendant, but that the defendant's pricing makes it unprofitable for the plaintiff to continue to produce. (Para. 469)⁸⁴

On the facts of the case, the District Court found that the discounted prices exceeded Abbott's average variable costs and that it was not unprofitable for Ortho to continue supplying the medical tests. It therefore dismissed the plaintiff's case.

In *LePage's*, an Appeals Court condemned 3M's bundled rebates under the monopolisation provisions. 3M had been offering multi-product rebates to large stationery, drugstore and grocery chains such as Staples, Office Depot, CVS, Walmart and K-Mart.^{85,86} 3M was found to have substantial market power in the US market for transparent tape, through its Scotch product. *LePage's*, by contrast, was mainly active in the supply of *private label* transparent tape, with a share of this segment of 88 per cent in 1992, falling to 67 per cent by the time of the trial. *LePage's* argued that 3M's practices caused it (*LePage's*) to forgo sales and to see its scale economies reduced (although no quantitative exercise was carried out), hindering its ability to compete going forward. The Appeals Court was receptive to this argument: according to the Appeals Court, 3M's bundled rebate programme harmed *LePage's*, because *LePage's* smaller volumes and narrower scope of products made it difficult for *LePage's* to compete over the products that both firms supplied; and this amounted to unlawful conduct under the monopolisation provisions.⁸⁷

⁸⁴ This is sometimes referred to as the '*Ortho* test'. See Section 4.5 for a discussion of price-cost tests in the context of bundled rebates.

⁸⁵ Between 1993 and 1994, 3M applied the Executive Growth Fund ('EGF') rebate program. This set individual growth targets for a small number of companies, with discounts ranging between 0.2 and 1.25 per cent of total sales conditional on achieving targets over three or more product categories among the following six: stationery, home improvement (for example, cleaning products), leisure (for example, audio/video products), home care, health care and retail auto products. From 1995 (and through 1999, the end of the period over which the infringement was found) 3M replaced the EGF with the Partnership Growth Fund ('PGF') rebate program. This offered rebates of up to 2 per cent of total sales, conditional on achieving targets over a minimum of two out of the six product categories listed above. The higher the number of product categories where the sales target was met, the higher the rebate. The targets in the PGF were not customer-specific, they applied to any customer.

⁸⁶ There were also allegations that 3M had made a few offers to specific retailers aimed at diverting business away from 3M's competitors towards 3M, akin to exclusive dealing clauses. The Court interpreted these as reinforcing the anti-competitive effect of the bundled rebates.

⁸⁷ The Supreme Court declined to review the case, upon the recommendations filed by the Department of Justice and Federal Trade Commission. These two bodies had urged

The results of a price-cost test in this case appeared to have little (if any) evidentiary value. Parties did not contest that 3M was more efficient than LePage's, nor that 3M priced its products above its costs, however calculated. In fact, as Rubinfeld (2005) noted, no evidence was presented, nor did LePage's allege, that 3M sold its tape at prices below cost, *even if the entire rebate across all products was attributed exclusively to transparent tape*. Yet this evidence did not sway the majority opinion of the Court.⁸⁸

In any event, the judgment was far from unanimous. As a dissenting judge on this case (joined by two further judges) wrote:

I conclude that as a matter of law 3M did not violate [antitrust laws] by reason of its bundled rebates even though its practices harmed its competitors. The majority decision which upholds the contrary verdict risks curtailing price competition and a method of pricing beneficial to customers because the bundled rebates effectively lowered their costs. I regard this result as a significant mistake which cannot be justified by a fear that somehow 3M will raise prices unreasonably later. (pp. 63–4)

Other Appeals Court's judges have been similarly critical of the majority opinion in *LePage's*⁸⁹ and so have several commentators. For example, Rubinfeld (2005) also noted that there was no evidence that the bundled rebates entailed a profit sacrifice for 3M; that is, LePage's had not shown that 3M could have achieved its legitimate business goals through a different, less costly commercial strategy (for example, single-product rebates). Rubinfeld (2005) also posited that if 3M really wanted to exclude LePage's, it could have done so – lawfully – by simply pricing its private label transparent tape very aggressively: given it was more efficient than LePage's, 3M would not have needed to price below its own cost in order to foreclose LePage's; and the *Brooke* standards (see Chapter 1) would have suggested that this conduct would be lawful.

the Supreme Court to wait until the economics literature had satisfactorily clarified the conditions that make bundled rebates anti-competitive.

⁸⁸ Further, it is worth noting that customers were not rewarded for specifically increasing their purchases of 3M's transparent tape: they were rewarded for increasing their purchases of stationery products as a whole (along with products in the other product categories), where stationery products would include any stationery product, *including non-3M transparent tape*. As such, a plausible hypothesis for 3M's rebate programme was that 3M was seeking to increase its business with retailers across a range of products, not simply stationery products, and *a fortiori* not just transparent tape.

⁸⁹ The Appeals Court judges in *PeaceHealth* (which we review below) went as far as stating that the LePage's standard could protect a less efficient competitor at the expense of consumer welfare (see para. 899 of the *PeaceHealth* judgment).

A few years later, in *PeaceHealth*, the Appeals Court (Ninth Circuit) explicitly rejected the Third Circuit's approach in *LePage's*.⁹⁰

McKenzie, a hospital services provider in Lane County, Oregon, sued PeaceHealth, its only competitor, for allegedly trying to monopolise the primary and secondary care markets (where PeaceHealth had a 75 per cent market share). The alleged anti-competitive conduct entailed PeaceHealth offering insurers discounts on tertiary care services (that is, very specialised and advanced treatments, where PeaceHealth was the only provider) of up to 40 per cent, provided PeaceHealth was chosen as the sole preferred provider of primary, secondary and tertiary services. This, McKenzie argued, made it impossible for it to compete against PeaceHealth, despite McKenzie having lower costs of providing primary and secondary care services.

The Court's judgment stood in stark contrast to that of *LePage's*. The Court in *PeaceHealth*, largely following *Ortho Diagnostic*, held that the relevant test in the context of a multi-product bundling (or bundled rebates) case was the 'discount attribution standard':

Under this standard, the full amount of the discounts given by the defendant on the bundle are allocated to the competitive product or products. If the resulting price of the competitive product or products is below the defendant's incremental cost to produce them, [...] the bundled discount [may be found] exclusionary for the purpose of [the monopolisation provisions]. This standard makes the defendant's bundled discounts legal unless the discounts have the potential to exclude a hypothetical equally efficient producer of the competitive product. (Para. 906)

The Court further added that the correct cost benchmark to implement the standard just set out was average variable cost. While the Court endorsed a price-cost test, it also set a clear distinction between the assessment of single-product and of multi-product rebates: specifically, it openly rejected the requirement for a plaintiff to show in a claim of anti-competitive multi-product rebates that the defendant incurred a loss and that there was a dangerous probability of recoupment of such loss; these are elements which are instead required for a successful claim of anti-competitive single-product rebates (given these claims are typically assessed under a standard predatory test, as discussed in Chapter 2).⁹¹

Ultimately, the Appeals Court therefore held that the District Court had erred in instructing the jury to follow a test based on *LePage's* (as opposed

⁹⁰ Here we also draw from Economides and Lianos (2009) and from White and Good (2009).

⁹¹ The Court reasoned that because a firm may fail to meet the discount attribution standard and yet not make a (financial) loss, it did 'not think it is analytically helpful to think in terms of recoupment of a loss that did not occur'. See footnote 21 of the judgment.

to the discount attribution standard) for their assessment of whether PeaceHealth's bundled rebates were unlawful under the monopolisation provisions.

4.6.2 European Union

Tying Just as we observed in the context of exclusive dealing in Chapter 3, there is also scant jurisprudence in Europe on tying and bundled rebates. This may be due to the fairly formalistic approach that the Commission and the Courts followed over early landmark cases. In summary, this approach was such that tying behaviour would be abusive when three conditions were satisfied: first, the alleged infringer had a dominant position in the supply of the tying product; second, the tying and tied products were distinct products; and, third, there was an element of coercion of customers, who had to buy the two products together. In this respect, the early EU case-law on tying was not too dissimilar from the early US case-law, both largely dealing with the tying of products in a 'primary market – after-market' relationship (see our discussion in Section 4.3.6). One can appreciate this in *Hilti* and *Tetra Pak II*. *EFIM* (where a 'primary market – after-market' relationship also applied) is instead a more recent case, and one where an economics-based approach was endorsed by the Courts. We discuss in turn each of these cases.⁹²

The first European landmark case on tying was *Hilti*. This arose in the market for nail guns (where Hilti had a market share of approximately 55 per cent) and its complementary markets for cartridges, cartridge strips and nails (where the company had around 70–80 per cent of the market for Hilti-compatible nails). Among a number of commercial practices scrutinised by the Commission, Hilti engaged in the tying of cartridge strips to nails with a number of distributors and end-users and refused to supply Hilti-compatible cartridge strips to independent Hilti-compatible nail producers. Hilti defended its tying requirements on the grounds of safety issues, but the Commission rejected this justification. The Commission found Hilti's practices anti-competitive mainly on the basis that they foreclosed the nascent competition from independent producers of Hilti-compatible nails and that consumers were being 'exploited' by being denied choice when purchasing nails.⁹³ The Courts upheld the

⁹² We also direct the interested reader to Temple Lang (2011) for a review of several cases involving after-markets.

⁹³ The Commission referred to this notion of exploitation at para. 75 of its decision.

Commission's decision. The General Court agreed with the Commission on the finding that guns, cartridge strips and nails were distinct products and they did not constitute (as Hilti argued) a single 'powder-actuated fastening' (or 'PAF') system. In fact, according to the General Court, the Commission was correct in defining even narrower markets, namely for Hilti-compatible cartridge strips and for Hilti-compatible nails. The General Court also endorsed the Commission's reasoning on Hilti's practices foreclosing competition in the nail market as well as the Commission's rejection of Hilti's arguments in relation to safety issues.⁹⁴ Here, we find that three aspects are worth discussing, in light of our previous review of after-markets in Section 4.3.6. First, the Commission and the Courts probably dismissed the existence of a PAF system (or more generally of competition between systems) a bit too hastily and using relatively weak arguments (namely that there existed some suppliers manufacturing nails and cartridge strips but not nail guns). Second, there was no assessment of competition in the primary market (nail guns).⁹⁵ Third, the Commission's assessment also neglected to consider how end-users actually behaved (including any switching behaviour between PAF systems or beyond) or what their expectations were at the time of buying a nail gun, for example. The Commission had noted in passing that end-users of these products were mainly professional users in the construction industry (rather than do-it-yourself enthusiasts, say). As such, one may have expected them to have a good sense of their requirements, and quite possibly regularly engage in whole-life costing (that is, in an estimation of the total costs of a given option, including those of consumables over time) and in a qualitative comparison of different alternatives available in the marketplace.

In *Tetra Pak II*, the European Commission scrutinised a number of potentially exclusionary practices carried out by the eponymous firm. In this regard, the Commission took issue with Tetra Pak tying (through an exclusive supply obligation) the sale of liquid packaging machines to the cartons it sold to the food industry, as well as to machine repair and

⁹⁴ The Court of Justice also upheld, on appeal.

⁹⁵ This was the key message in Shapiro (1995), as discussed in Section 4.3.6: a sufficient degree of competition in the primary market may dissipate any rents in the after-market(s), given that these would be competed away as part of the competitive process in the primary market; so tying could not be a way to set prices that exploit end-users or that exclude competitors. By contrast, in relation to nail guns, the Commission's assessment of the competitive process was almost completely absent and almost solely relied on a cursory reference to market shares and patents.

maintenance services and spare parts. The Commission found that Tetra Pak held a ‘virtual monopoly [in the] field of the aseptic packaging of long-life products’ (para. A.I.(1)), with a 92 per cent share at EU level in the aseptic packaging machines market and an 89 per cent share in the aseptic cartons market.⁹⁶ The Commission also found that Tetra Pak had a leading position in the respective non-aseptic markets (that is, in relation to the packaging of other fresh liquid foods), with a 52 per cent share of the non-aseptic packaging machines market and a 48 per cent share of the non-aseptic cartons market.⁹⁷ Tetra Pak justified its tying on efficiency grounds, arguing this drove economies of scale and cost savings on raw materials and distribution. Tetra Pak further argued that its behaviour should be justified (i) on health and safety grounds, (ii) on reputation grounds (that is, to protect the firm’s reputation) and (iii) on liability grounds (that is, to guarantee customers a single source of liability in the case of product faults). The Commission rejected these claims. It reasoned that Tetra Pak’s ‘virtual monopoly’ in the aseptic markets, combined with a leading position in the non-aseptic markets, made it ‘the inevitable supplier for a majority of users’ (para. 105). According to the Commission, Tetra Pak managed to impose on such users ‘certain contractual obligations [including exclusive supply obligations, resulting in *de facto* tying] aimed essentially at binding them to the group and preventing any trade in its products’ (para. 105), eliminating both intrabrand and interbrand competition. Tetra Pak’s appeals to both Courts were unsuccessful. The General Court found that the exclusive supply (tying) obligations were part of a broader strategy employed by Tetra Pak, supported by a wider range of contractual clauses:

[T]he combined effect of the other 24 contractual clauses at issue [...] was an overall strategy aiming to make the customer totally dependant on Tetra Pak for the entire life of the machine once purchased or leased, thereby excluding in particular any possibility of competition at the level both of cartons and of associated products. Their effect on competition must therefore be considered in conjunction with [other] clauses [applied by Tetra Pak] which were intended to make the market in cartons wholly dependent on that in machines and which reinforced and completed the elimination of that market. Moreover, those other clauses could be considered as abusive in themselves since their object was in particular, depending on the clause, to make the sale of machines and cartons subject to accepting additional

⁹⁶ ‘Aseptic’ referred to the incorporation of technology for the sterilisation of cartons and the packaging in those cartons, under aseptic conditions, of ultra-high temperature (‘UHT’)-treated liquid foods.

⁹⁷ Further, the Commission noted that Tetra Pak owned many of the patents on the technology used in the relevant production processes.

services of a different type, such as maintenance and repair and the provision of spare parts [...]. (Para. 135)

In *EFIM*, a number of independent suppliers of generic ink cartridges (grouped into the European Federation of Ink and Ink Cartridge Manufacturers, 'EFIM') lodged a complaint to the European Commission on the basis that they were denied access to intellectual property rights by four manufacturers of printers (collectively, original equipment manufacturers or 'OEMs'). EFIM claimed that without access to such intellectual property rights, independent suppliers could not effectively compete with the OEMs in relation to ink cartridges, an after-market. The Commission dismissed EFIM's complaint mainly on the basis that it found it questionable whether a firm could be dominant in an after-market if it was not dominant in the primary market. In other words, intense competition in the (primary) market for printers would prevent the exercise of market power in the supply of any brand-compatible ink cartridges. Referring to an earlier rejection of a similar complaint, the Commission noted that dominance in an after-market could be ruled out where: (i) a customer can make an informed choice, including in relation to whole-life costs; and (ii) is likely to make such an informed choice accordingly; (iii) should a supplier seek to exploit customers in one specific after-market, a sufficient number of customers would adapt their purchasing behaviour in the primary market; and (iv) they would adapt such behaviour within a reasonable timeframe. The General Court upheld the Commission's decision to reject the complaint, endorsing the Commission's economic thinking. And so did the Court of Justice, which likewise endorsed it, including insofar as the four above-mentioned conditions were concerned. The Court of Justice also inferred from these conditions that dominance in the after-market can be ruled out in the presence of competition in the primary market and of a close linkage between the primary market and the after-market (see para. 37 of the judgment).

Few antitrust cases have received as much media attention in Europe as *Microsoft*, which led to a record fine (at the time) of €497m. Here, we discuss the tying aspects of this case, between the Windows' operating system and Windows Media Player.⁹⁸ The Commission found Microsoft to be dominant in the market for PC operating systems (with a market share in excess of 90 per cent). Windows Media Player was offered to customers

⁹⁸ Microsoft was also found to have unlawfully refused to supply interoperability information for work group server operating systems; we discuss this part of the case in Chapter 5.

(free of charge) already embedded in the Windows operating system and it could not be uninstalled. The Commission rejected Microsoft's objective justifications centred on technical efficiencies from tying and concluded that early antitrust intervention in that rapidly evolving market was warranted to avoid a tipping outcome, which would stifle future competition in the market for media players, as well as in the adjacent ones for software and content. The General Court agreed with the Commission's conclusions, endorsing the approach to the assessment of tying through an evaluation of: dominance in the tying market; whether the markets for the tying and the tied products were separate; whether the conduct in question implied that the tying product (Windows' operating system) could not be obtained without the tied product (Windows Media Player); whether tying led to anti-competitive foreclosure in the tied market; and whether there were objective justifications capable of offsetting any anti-competitive effects identified. The Court also agreed with the Commission's assessment that the tying behaviour had a spillover effect on competition in related products, such as media encoding and management software, but also in client PC operating systems for which media players compatible with quality content are an important application.⁹⁹ We discuss this case in more detail in Section 4.7, drawing a comparison, from an economic perspective, with Microsoft's tying case in the US (which concerned web browsers).

In *Microsoft (tying)*, the European Commission brought a similar case against Microsoft, which allegedly engaged in the anti-competitive tying of its web browser (Internet Explorer). Specifically, as its main theory of harm, the Commission considered that the tying of Internet Explorer to Windows may have had the potential to foreclose the market for web browsers and that the tying gave Internet Explorer an artificial distribution advantage that other web browsers were unable to match. As additional theories of harm, the Commission also considered that Microsoft's tying may have (i) limited web development (due to the existence of network effects which would have reduced the incentives for content providers and software developers to support web browsers other than Microsoft's Internet Explorer) and (ii) reinforced Microsoft's position in the client PC operating system market. On that occasion, though, differently

⁹⁹ More generally, some commentators (for example, Ahlborn and Evans, 2009, and Ezrachi, 2016) note that the General Court largely ignored the issue of whether and in what form consumer harm arose and in fact simply presumed it, based on the facts of the case. In this respect, as Ahlborn and Evans (2009) note, the General Court appeared to reject the Commission's 'invitation' to move from a *per se* rule towards an effects-based approach to the assessment of tying conduct by a dominant firm.

from the Windows Media Player case, the Commission did not reach an infringement decision: in December 2009 Microsoft offered a set of commitments, including providing new users or users who at the time had Microsoft Internet Explorer as their default browser with a ‘choice screen’. This would feature the 12 most popular web browsers at a given point in time, which a user could choose from.¹⁰⁰

In *Rio Tinto Alcan*, the Commission followed the test for anti-competitive tying that was endorsed by the General Court in *Microsoft*. The Commission was concerned with the practice by Rio Tinto Alcan (‘RTA’) of tying the licensing of its aluminium smelting technology to the supply of pot tending assemblies (‘PTAs’, that is, large handling cranes used in smelter plants). The Commission found *prima facie* evidence of an abuse of dominance, but agreed in December 2012 to terminate its proceedings following RTA’s undertaking to discontinue its contractual tying. We discuss this case in more detail in Section 4.7.

At the time of writing, there are a number of ongoing investigations by the European Commission involving Google. In the context of practices that are relevant to the discussion in this chapter, there are open proceedings in *Google (Android)*.¹⁰¹ The Commission noted that about 80 per cent of smart mobile devices in Europe run on Android, the mobile operating system developed by Google, which Google licenses to device manufacturers.¹⁰² In *Google (Android)*, the Commission has *inter alia* been taking issue with Google’s practices to require devices manufacturers: (i) to pre-install Google Search on their devices and set it as the default search service as a condition to license Google’s Play Store app;¹⁰³ and/or (ii) to pre-install Google’s Chrome browser on their devices, as a condition

¹⁰⁰ Through another commitment, original equipment manufacturers would be free to pre-install any web browser(s) of their choice on PCs they shipped and set it as the default web browser.

¹⁰¹ In what follows, we draw on the European Commission’s press release IP/16/1492 and on its fact sheet MEMO/16/1484, both dated 20 April 2016.

¹⁰² The Commission noted that Google’s market shares in the European Economic Area for licensable mobile operating systems exceeded 90 per cent; and that Android was used on ‘virtually all smartphones and tablets in the lower price range, which are bought by the majority of customers’ (see European Commission, MEMO/16/1484, 20 April 2016). The Commission also pointed to network effects acting as entry barrier protecting Google’s position (arguing that the more consumers adopt an operating system, the more developers write apps for that system). The Commission also highlighted the existence of significant switching costs, in the form of users losing their apps, data and contacts if they wished to switch to other operating systems.

¹⁰³ The Commission noted that Google’s Play Store app accounted for more than 90 per cent of apps downloaded on Android devices in the European Economic Area.

to license Google's Play Store app or Google Search.¹⁰⁴ The Commission was concerned that the practices challenged in *Google (Android)* may (i) strengthen Google's dominant position in general internet search services; (ii) affect the ability of competing mobile browsers to compete with Google Chrome; (iii) hinder the development of operating systems based on the Android open source code and the opportunities they would offer for the development of new apps and services; and (iv) ultimately harm consumers as they are not given as wide a choice as possible and due to the stifling of innovation.

Bundled rebates EU jurisprudence does not formally distinguish between single- and multi-product (or bundled) rebates. We refer the reader to Chapter 2 for further discussion of the case-law on single-product rebates and summarise next the most notable cases on bundled rebates.

In *Hoffmann-La Roche*, probably the most notable EU case on bundled rebates, the Commission ruled against Hoffmann-La Roche's use of contractual clauses in the early 1970s, which imposed single branding or offered loyalty rebates to the purchasers of its synthesised vitamins (mostly sold to customers in the pharmaceutical, food and animal feed industries). The rebates (of around 1 to 5 per cent, but in some instances of up to 20 per cent) were offered upon the condition that most or all vitamin requirements by a given customer came from Hoffmann-La Roche (which in addition offered price-matching schemes through an 'English clause'). The Commission found the effect of these rebates particularly strong due to the fact that customers bought several vitamin categories from their suppliers. There was also a (near) single-branding requirement for each category: a firm failing to source 90 per cent of its vitamin A from Hoffmann-La Roche would also lose the loyalty rebates on vitamin C, say. In sum, the Commission found an abuse of a dominant position due to Hoffmann-La Roche's behaviour hampering customer choice and restricting competition between bulk vitamin manufacturers in the EU. The Court of Justice agreed with the Commission's assessment (with the exception of the dominance assessment in one of the relevant product

¹⁰⁴ The *Google (Android)* case involves other practices that are under scrutiny at the time of writing – see Chapter 3 for brief remarks. In Chapters 3 and 5 we also refer to two further, separate, investigations by the Commission involving Google. One of these, *Google (comparison shopping)*, may potentially be assessed as a tying case: under this interpretation, Google would be tying its Google shopping service pages to the pages featuring the results of Google's organic search. Since the case may also be assessed under a vertical foreclosure framework (whereby Google would favour its comparison shopping service in its search result pages) we return to it in Chapter 5.

markets). As Whish and Bailey (2015: 723–4) note, the test used by the Court of Justice was largely based on a treatment of loyalty rebates of this kind as single-branding agreements, making them unlawful. The Court’s reasoning is summarised as follows:

Obligations [...] to obtain supplies exclusively from a particular undertaking, whether or not they are in consideration of rebates or of the granting of fidelity rebates intended to give the purchaser an incentive to obtain his supplies exclusively from the undertaking in a dominant position, are incompatible with the objective of undistorted competition within the Common Market, because [...] they are not based on an economic transaction which justifies this burden or benefit but are designed to deprive the purchaser of or restrict his possible choices of sources of supply and to deny other producers access to the market. (Para. 90)

This appears to us as a very formalistic approach, as it seems to infer that the exclusivity obligation is by its nature restricting the choice of buyers and excluding rivals. It is also an approach that did not include a price-cost test nor an assessment of the effects of the conduct (although admittedly this was a very early case in the jurisprudence).

A more recent case on bundled rebates was *Coca Cola*, which concluded in 2005 with the soft drinks company offering commitments to the European Commission, thus avoiding a full investigation. In addition to exclusivity clauses, growth and target rebates, constraints on shelf-space use and other contract clauses, the Commission appeared to be concerned with Coca Cola’s use of multi-product discounts in such a way as to promote the sale of its weaker brands in an anti-competitive fashion (the short decision, only one-page long, does not allow for a more detailed assessment of the facts).

4.7 Cases

In this section, we discuss a number of cases from various jurisdictions. These cases may lend themselves relatively well to an evaluation of whether the facts in those cases (as reported in decisions and judgments) could be seen as consistent with the key aspects of the economic models reviewed earlier in this chapter. We discuss three tying cases first, before turning to one bundled rebates case.

4.7.1 Microsoft Internet Explorer – US

We summarised the tying part of the US *Microsoft* case in our earlier review of the US case-law on tying. Here, we investigate the economics of this case

in some detail.¹⁰⁵ The tying model put forward by Carlton and Waldman (2002) – which we presented in Section 4.3.4 and in technical Sections 4.4.3 and 4.4.4 – was in fact motivated by the *Microsoft* case.

The case focused on ‘applications barriers to entry’, namely the difficulty that Microsoft’s competing operating systems (or more generally platforms) and potential entrants allegedly faced due to the large number of applications that were compatible with the Windows’ operating system but not with others.

Gilbert and Katz (2001) discuss the case in detail, and provide an interpretation of the economics of the case that dovetails with Carlton’s and Waldman’s model. In particular, they point out that the case brought by the US Department of Justice focused on the role of the Java technology as a middleware. Java supported a programming language that enabled applications (such as word processing and spreadsheets) to run independently of operating systems; Java was also used by Netscape (including by its Navigator web browser). Hence, the threat to Microsoft, the argument went, was that if applications using this programming language reached a critical mass of users, programmers would have probably had sufficient incentives to write an increasing number of different applications that would run on it. This would have in turn increased the likelihood that the Windows operating system would face a fierce competing platform (whether or not an actual operating system). Gilbert and Katz (2001) even suggest that there was a possibility that Netscape Navigator itself could eventually establish itself as a competing platform to the Windows operating system.

Therefore, against this background, the argument was that Microsoft’s tying of Internet Explorer (the web browser) to Windows (the operating system) was a way to marginalise Netscape Navigator due to its role of promoting the adoption of Java technology by users and programmers alike.¹⁰⁶ To be clear, therefore, this theory of harm was *not* around any leveraging of market power from the supply of operating systems to that of web browsers, but rather on the defence of Windows’ market power in the supply of operating systems.

Microsoft, from its side, retorted that bundling Internet Explorer to Windows generated cost savings (via lower distribution and transaction costs); and that competition took place between platforms, where a

¹⁰⁵ For a thorough review of this case, the reader is also directed to Motta (2004: section 7.5).

¹⁰⁶ The Appeals Court also referred to evidence of intent of Microsoft’s exclusionary behaviour (see for example paras 154–61 of the judgment).

platform comprised an operating system together with various applications sought after by end-users. Hence, it argued, it made little sense to study competition within the supply of only one component of such platforms. It also argued that the bundling of the operating system and the web browser constituted a product improvement.

While such arguments may have provided a *legal* rebuttal to the charges received (which we are not best placed to comment upon), we note that these arguments would not necessarily address the rationale underpinning the theory of harm set out above, which we discuss next in more detail.

The facts in *Microsoft* appear to be consistent with the network effects version of Carlton and Waldman's model (which we discussed in more detail in technical Section 4.4.4). Recall that a crucial factor for this model to be used as a possible theory of harm is a clear linkage between one market which is currently open to competition by a rival and one which may be open to competition by this rival in the future *only so long as sufficient scale is reached in the first market*. Building on the arguments put forward by the US Department of Justice, and consistent with Carlton and Waldman (2002) and Gilbert and Katz (2001), it would appear that the network effects would arise in the form of the success of the Java technology depending on a sufficient uptake of it by users and programmers; and that Netscape Navigator was a key component of such uptake. Hence, achieving a sufficient scale in the web browser market was seen as a critical condition to generate sufficient interest in the Java technology to support applications that would not run on the Windows operating system. In this sense, the achievement of a sufficient scale by Netscape Navigator in the web browser market was seen as instrumental to enable a challenge to the Windows operating system at a later stage. This was the market to be protected through the tying, according to the spirit of Carlton and Waldman's model.¹⁰⁷

In sum, the main facts in the US *Microsoft* case appear to be consistent with the theory of harm put forward in Carlton and Waldman (2002), whereby tying (building on scale and/or scope economies, in this case on the demand-side through network effects) does not aim to 'leverage' any

¹⁰⁷ By contrast, we note that the tying of Windows and Internet Explorer was not technical and could in fact be reversed. However, it is important to note that for this variant of the Carlton and Waldman (2002) model (unlike the supply-side scale economies variant) the credibility of the commitment to tie is *not* a requirement.

market power to another market, but rather aims to protect a given market position in the primary market.

4.7.2 Microsoft Windows Media Player – EU

Prima facie, the European tying case against Microsoft in relation to the Windows Media Player (briefly summarised in Section 4.6) may seem quite similar to the US case (concerning web browsers) we have just reviewed. In fact, the underlying facts are fairly different between the two cases (based on the information set out in the respective decisions and judgments), and the Carlton and Waldman (2002) model reviewed extensively in this chapter is an unlikely candidate to describe the economics of the EU case.

As we have just discussed, in the US case, Microsoft was accused of using its market power in operating systems (the primary market) to foreclose competition in web browsers (a complementary market) with the ultimate aim of hindering the development of an alternative technology (Java) that would have competed against Microsoft *in the primary market*. In the EU case, the first leg of the argument was similar, namely Microsoft was accused of using its market power in operating systems to foreclose competition in the complementary market (media players). This complementary market was itself a two-sided market: limited adoption of a competing media player would mean that programme developers may see their incentives to develop specific computer codes and applications tinkered with, thus perpetuating Windows Media Player's strong position in the market (and possibly Microsoft's position in adjacent markets). However, there was only a passing mention to the risk that competition in the primary market (operating systems) may in turn be affected by the degree of competition or by the technologies adopted in the market for media players (see paras 973–4 of the Commission's decision). As such, while the Carlton and Waldman (2002)'s model may represent a reasonable approximation of the facts of the US case, this would not appear to be substantiated for the EU one.

The Commission structured its assessment focusing on the following elements: Microsoft's dominance in the tying market (the operating system); the tying and the tied products being distinct; coercion, namely Microsoft not giving consumers the choice of purchasing the operating system without the media player; and foreclosure effects.¹⁰⁸ To assess this

¹⁰⁸ Ahlborn and Evans (2009) note that while the Commission at least sought to assess any likely or actual harm to competition, the General Court found that the circumstances of

last element, the Commission largely relied on the ubiquity of Windows Media Player among consumers. It then set out its concerns about the future incentives that firms would have in adjacent markets (such as media encoding and management software), once the media player market ultimately tipped towards Microsoft. The following excerpts from the decision clarify the Commission's concerns:

In a nutshell, tying WMP [Windows Media Player] with the dominant Windows makes WMP the platform of choice for complementary content and applications which in turn risks foreclosing competition in the market for media players. This has spillover effects on competition in related products such as media encoding and management software (often server-side), but also in client PC operating systems for which media players compatible with quality content are an important application. Microsoft's tying practice creates a serious risk of foreclosing competition and stifling innovation. [...] Through tying WMP with Windows, Microsoft ensures that WMP is as ubiquitous on PCs worldwide as Windows is. No other distribution mechanism or combination of distribution mechanisms attains this universal distribution. Microsoft obviously controls this mode of distribution and (no longer) affords competing media player vendors access to it. (Paras 842–4, footnotes omitted) [...]

It should also be highlighted that Microsoft has a clear incentive to attain a strong position in the media player market due to the fact that it can use the ubiquity of WMP on Windows PCs as an argument to propagate its proprietary media formats and technologies at the server software side, in its relationship with content developers (leveraging effect). The media player market is, in fact, a strategic gateway to a range of related markets, on some of which high revenues can be earned. As such [...] gaining a pre-eminent position in the media player market will provide Microsoft with a significant advantage in other business areas such as those for content encoding software, format licensing, wireless information device software, [digital rights management] solutions and online music delivery. (Para. 975, footnotes omitted)

The Commission found that it was quite costly for content and software developers to design products for more than one platform, because of different technical specifications.¹⁰⁹ This compounded the Commission's foreclosure concern.

Kühn et al. (2005) also interpret this case in a similar fashion to that of the Commission:

the case (in particular, the inability of Microsoft's competitors to replicate the bundle) were sufficient to establish harm and this was an irrefutable presumption.

¹⁰⁹ A survey of software developers prepared for the Commission indicated that the cost of a second version for a different platform would entail, on average, an additional expenditure equal to 58 per cent of the original version.

It is clear, then, that bundling by a firm with dominance in the PC operating system market can in theory completely tip the market for encoding software to itself by technically bundling the operating system with its media format decoding software. In that case it achieves ubiquity on all PCs while the delivery of decoders for other formats will always be imperfect and is empirically of questionable efficacy. This is therefore another example of how bundling today can completely change competition in the market in the future. It should be noted that the dominant position in the operating system market is critical here for the anticompetitive effect of bundling.

These authors further note that the potential consumer harm could be very large, in terms of foregone quality improvement and innovation:¹¹⁰

Media formats and encoding technologies are at the forefront of research for Internet technologies. There is great value from finding efficient encoding methods to increase the speed of delivery of content.

For its part, Microsoft retorted, *inter alia*, that the media player was an integral part of the operating system, so that tying was not a meaningful concept in that context, but the Commission disagreed with this argument.¹¹¹

In sum, we find that the Commission could have spelled out its theory of harm in a clearer way. The key channel of potential harm the Commission explored was in relation to foreclosure to competition *outside* the tying-tied product pair, in relation to media encoding and management software; but the Commission also alluded to harm to competition in the tied market (media players) and (very briefly, and without much substantiation) to the possibility of harm to competition ultimately in the tying market (operating systems). In the decision, however, the distinction between these channels was not particularly clear, and the anti-competitive effects alluded to were to some extent speculative.

With hindsight, we find that the model proposed by Carlton and Waldman (2012) – reviewed in Section 4.3.2.3 – may support a theory of harm in line with the concerns set out by the Commission: namely

¹¹⁰ Moreover, these authors stress that the elimination of effective competition in media players (leading to lower quality) would not necessarily require the competitors to exit the market; stopping investing in new technologies would be sufficient for the elimination of effective competition.

¹¹¹ Specifically, the Commission made a number of points, including: there was demand for stand-alone media players, distinguishable from demand for PC operating systems; a number of vendors developed and supplied media players on a stand-alone basis; Microsoft itself developed and distributed versions of its media player for other PC operating systems; and Microsoft promoted its media player in direct competition with third party media players.

one based on tying as a way to foreclose competition in the development and supply of complementary products that have yet to come to market. As we discussed when reviewing this model, the standard Chicago School argument would be that the incumbent or dominant firm would be able and have an incentive to appropriate any efficiencies of the rival (or the rents that the rival may generate in the future through the development of new or better complementary products) contractually or through the pricing of the primary product. However, under some conditions, the incumbent may be unable to appropriate such efficiencies (or rents) and may therefore have an incentive to resort to tying or bundling thus foreclosing the rival. An important assumption for this model to apply is that a rival is *de facto* foreclosed as a result of the tying or bundling, for example because all or a large proportion of consumers would not buy or stock multiple versions of a complementary product (say, several media players). This last aspect was fairly heavily debated in *Microsoft*.

4.7.3 Rio Tinto Alcan

In *Rio Tinto Alcan*, the European Commission was concerned with some contractual practices of Rio Tinto Alcan ('RTA') in relation to the licensing of aluminium smelting technology and to the supply of pot tending assemblies (or 'PTAs', that is, large handling cranes that perform various operations in aluminium smelting plants and without which a plant could not produce aluminium, according to the Commission). In particular, the Commission was concerned about RTA's practice of tying its smelting technology licences to the exclusive purchase of PTAs from RTA. The Commission reasoned that this may lead to negative effects on innovation and prices (see below for further discussion) and result in anti-competitive foreclosure in the PTA market.

To address the Commission's concerns, RTA offered to remove from its technology licences (for a period of five years) the clauses requiring licensees of its aluminium smelting technology to purchase PTAs from RTA and to replace them with an obligation to purchase PTAs some from 'pre-qualified suppliers'. The Commission was satisfied that this commitment would address the concerns set out and thus closed the investigation without an infringement finding.

In what follows, we first briefly summarise the Commission's competitive assessment, before setting out a possible interpretation of this case from an economic perspective.

In relation to the ‘tying market’ (licensed aluminium smelting technology) the Commission ruled out the effective substitutability of licensed technologies with in-house technologies. Neglecting in-house technologies, the Commission calculated that RTA carried out between 70 and 80 per cent of all licensed smelter projects implemented from 1995 to 2006 outside China (which was considered a distinct relevant geographic market).¹¹² Besides a handful of Chinese suppliers, the Commission only identified two other suppliers active in this industry in addition to RTA.

The Commission also noted the presence of significant barriers to entry, due to a number of factors, including: significant scale economies, the need for significant levels of R&D and the need to build reputation around a technology’s reliability. As the Commission put it,

[t]he development of a licensable and competitive technology requires even more intensive R&D efforts and larger teams than those required for the operation of a smelter. To recover investment from technology development, large production capacities and a sizeable development team are needed, ideally combined with a large customer base that would allow for recovery of capital expenditures and R&D investments. (Para. 45 of the decision)

The Commission believed that there was no significant countervailing bargaining power due to the presence of numerous buyers, each with a relatively small proportion of the market for the production of aluminium.

As for the ‘tied market’ (the supply of PTAs) the Commission found that it too was characterised by high barriers to entry. As the Commission put it,

[t]he preliminary results of the Commission’s investigation indicate the existence of high barriers to entry into the PTA market: economies of scale and scope and the need for continuous investment into product-specific R&D as well as project-related investment (design and adaptation of PTAs to a specific smelter). (Para. 70 of the decision)

The Commission identified only four manufacturers of PTAs in the world, two of which were Chinese and almost exclusively supplied the Chinese market. The other two were RTA (through a subsidiary) and Réel (the company that originally raised concerns to the Commission about

¹¹² The Commission provided different market share estimates based on different parameters, such as installed capacities by technology, sales of primary aluminium and royalties fees paid for licensing of smelting technologies. The range of RTA’s market share was very wide depending on the parameter chosen, going from [40–50] per cent to [80–90] per cent.

RTA's conduct). The Commission found that RTA supplied 50–60 per cent of the global market for PTAs (excluding China).¹¹³

The Commission considered that the aluminium smelting technology licensing and the PTAs were distinct products, pointing out that the customers it sought evidence from stated that they would have preferred to source these inputs separately, absent the contractual tying restriction imposed by RTA. The Commission noted that RTA maintained significant market shares, despite charging more for its PTAs than Réel; it added that customers perceived RTA's PTAs as generally more expensive, albeit technically similar to Réel's PTAs.

The Commission dismissed the efficiencies flowing from tying that RTA had argued for. Specifically, it stated that 'any efficiency-enhancing tying must be driven by customer preferences for joint consumption' (para. 89), which was not the case in the industry at hand, according to the Commission, as noted above. Moreover, it had actually found that tied customers (licensees) were paying more for PTAs than non-tied customers, and it interpreted this as evidence going against the efficiency put forward by RTA.

The Commission did not set out a particular theory of harm, perhaps also because it limited itself to a preliminary assessment of the case, which was closed with a commitment decision.

However, the possible anti-competitive effects of tying may be in line with the exclusionary theories based on denial of scale economies and entry/expansion barriers that we have presented throughout the book. The Commission found that RTA enjoyed a strong incumbency advantage in the tied segment PTA market, since it had a strong base of large, tied customers, while Réel only had access to two large customers, neither of which had growth plans. And RTA reacted with aggressive pricing to Réel's entry into the contestable part of the relevant PTA market: the Commission identified some RTA internal documents that it claimed showed that RTA had strategically decided to 'reduce its gross margin [...] for PTA offers to non-tied customers with a view to foreclosing Réel from a sufficient customer base' (para. 74).

In other words, the little information available in the decision seems to point to a possible combination of tied sales (ranging between 20 and 40

¹¹³ As in the case of the licensing of aluminium smelting technology, the Commission found that institutional factors and the nature of competition in the supply of PTAs were very different between China and the rest of the world, and it therefore considered the relevant dimension of the geographic market to be worldwide excluding China.

per cent over 1999–2010) and aggressive pricing on non-tied customers in order to subtract scale to Réel, and marginalising or even leading it to exit the market.¹¹⁴ Indeed, the Commission found that the PTA revenues Réel was earning from its existing customers were insufficient for Réel's PTA division to remain viable.

The fact that RTA appeared to be following different commercial strategies with different customers (tying sales to some customers, and pricing PTAs aggressively to other customers which only purchased PTAs from RTA) may potentially be driven by the fact that some customers used in-house technologies for aluminium smelting and therefore did not require licensed technologies. While, as mentioned above, the Commission considered these technologies to be part of separate markets, we understand based on the information available that, in its assessment of the PTA market, the Commission considered all PTA sales, both to customers with in-house technologies and those with licensed technologies for aluminium smelting.

Ultimately, the Commission found that tying would have harmed competition, since the likely consequences of Réel's marginalisation and/or exit would have been: (a) customers being captive to RTA and (b) a reduction in the incentives to innovate in the PTA market.

With respect to the possible anti-competitive effects acting through reduced incentives to innovate of its rival, we note that – based on the information set out in its decision – there may also be some elements that could in principle be consistent with the Choi (2004)'s model presented in Section 4.3.3.3. In that model, tying acts as a device to commit to more aggressive pricing and a higher market share in the tied market. This raises the incentive to engage in R&D by the incumbent firm and reduces the incentive to invest in R&D for its rival, resulting in the leveraging of market power from the tying market to the tied market through the marginalisation (though not necessarily the complete exclusion) of the competitor in the tied market.

4.7.4 IDEXX

IDEXX Laboratories Limited ('IDEXX') was a company specialising in diagnostics and information technology solutions for animal health. In April 2010, the UK's Office of Fair Trading ('OFT', now Competition

¹¹⁴ Of course, aggressive pricing would be anti-competitive only to the extent that it is below an appropriate cost measure (see Chapter 1).

and Markets Authority) received a complaint that IDEXX was engaging in anti-competitive conduct. The OFT focused its investigation on three areas: (i) anti-competitive bundling of in-clinic (that is, on-site) analysers with external lab tests; (ii) anti-competitive bundling of specialist external lab tests with standard external lab tests; and (iii) predatory pricing in the in-clinic analysers market (where prices were allegedly below cost, conditional on a certain volume of consumables being purchased). We next summarise the OFT's assessment of each area in turn, including reviewing how the OFT applied a price-cost test, before providing some remarks.

Bundling of in-clinic analysers with external lab tests The OFT tested whether IDEXX was attempting to foreclose the external lab tests market, where it was not found dominant, by providing discounts, in the form of multi-product rebates, on its in-clinic analysers, where instead it was found dominant. In relation to the assessment of IDEXX's in the markets for in-clinic analysers, the OFT highlighted that IDEXX had a first-mover advantage, enjoyed brand reputation, faced customers who preferred sourcing their requirements from a single supplier, and that there were barriers to switching, barriers to entry, barriers to expansion and limited buyer power.

For the multi-product rebates to be likely to have a foreclosure effect, the OFT believed that three cumulative conditions would have to be met: a) the alleged infringer has a dominant position in the tying markets (that is, in-clinic analysers) and competitors in the tied market (that is, external lab services) would have to be unable to replicate IDEXX's bundle; b) a competitor as efficient as IDEXX in the tied market alone would be unable to match the effective incremental price (that is, the difference between the price of the bundled offer and the stand-alone price of an in-clinic analyser) in that market without making losses; and c) the extent of the alleged abuse would have had to be sufficient to restrict competition in the tied market.

The OFT assessed each condition and concluded as follows: a) IDEXX was dominant in each of the in-clinic analyser markets. The OFT added that it was likely that IDEXX's competitors would have found it difficult to replicate IDEXX's bundle due to: the presence of a single competitor which was active in the relevant markets, IDEXX's large market share, IDEXX's brand reputation, IDEXX's first-mover advantage and veterinary practices' general preference for a single supplier for their in-clinic analysers requirements. b) A competitor as efficient as IDEXX would be able to

match the effective incremental price in the tied market. The OFT reached this conclusion by carrying out a price-cost test: it concluded that the incremental price for external lab services was consistently above IDEXX's average avoidable cost ('AAC') of supplying those external lab services.¹¹⁵ Thus, a competitor as efficient as IDEXX in the external lab services market could not be foreclosed on the basis of IDEXX's multi-product rebates (we shall come back below to the choice of the AAC benchmark). c) In relation to the extent of the alleged abuse, the OFT found that the contracts involving multi-product rebates amounted to less than 5 per cent of the external lab services market. The OFT thus reasoned that even if competitors were foreclosed from this part of the market, IDEXX's conduct would have been unlikely to impair competition in the market as a whole.

It therefore dismissed this first area of potential concern.

Bundling of specialist external lab tests with standard external lab tests

The OFT also tested whether IDEXX's discounts on two specialist external lab tests (where it was found to be a monopolist) could foreclose the market for standard external lab tests (where it was not found dominant). IDEXX was offering such discounts conditional on a customer also purchasing standard external lab tests from IDEXX. Following the same analytical framework set out above, the OFT concluded that IDEXX's competitors could not replicate its bundle, on the basis that those specialist tests were proprietary IDEXX tests. The OFT also found that the bundle between some standard external lab tests (so-called 'Basic Profile') and one specialist external lab test led to an effective price of the standard external lab tests which was negative¹¹⁶ (and thus below AAC). However, it dismissed this area of potential concern on the basis that the extent of the tied product market affected by IDEXX's pricing practice was insufficient to impair competition in that market. In particular, the OFT found that in 2010 about 55–65 per cent of the customers who purchased one specialist test, and about 65–75 per cent who purchased the other, did so on a stand-alone basis, that is, not as a bundle. It also found that even if both specialist tests

¹¹⁵ The OFT carried out this test using a three-year horizon, on the basis that IDEXX told it that customers typically committed to buying consumables related to in-clinic analysers for this duration.

¹¹⁶ This was calculated to be -£1.50 and was computed as the difference between the price of a bundle between Basic Profile and the specialist external lab test (£33.50 + tax) and the price of the specialist external lab test alone (£35 + tax). The bundle between Basic Profile and the other specialist external lab test considered by the OFT led to an effective price of Basic Profile of £4.70, which may have been below cost according to the OFT.

had been systematically bundled with Basic Profile, only less than 2 per cent of the market for external lab tests would have been foreclosed, which was not deemed sufficient to raise competition concerns.

Predatory pricing of in-clinic analysers Finally, the OFT tested whether IDEXX's prices of in-clinic analysers (as well as of consumables associated with each type of in-clinic analyser) were predatory. It had received allegations that IDEXX at times gave away some in-clinic analysers free of charge. Once again, the OFT considered that the relevant benchmark was AAC: that is, absent any documentary evidence of predatory intent, only prices below AAC would be able to foreclose competitors. The OFT considered that there were likely to be 'systems markets', each comprising one type of in-clinic analyser (for example, biochemistry) and its associated consumables. Such consumables could therefore be considered after-market products, which we discussed in Section 4.3.6. The OFT also pointed out that, given that each contract was negotiated individually, the analysis should be carried out on a contract-by-contract basis, rather than on the basis of a generic average price. It noted that to the extent that customers (veterinary practices) engage in whole-life costing – that is, consider together the prices of a specific in-clinic analyser and of the related consumables over a period of time – both profits from analysers and consumables should be included in the analysis. Based on its analysis, the OFT found that IDEXX's prices (for in-clinic analysers and consumables) exceeded the AAC benchmark, even for contracts where an in-clinic analyser had been provided free of charge. It also considered the possibility that some veterinary practices did not engage in whole-life costing. In that case, the OFT reasoned that the relevant predation test should not take revenues from consumables into account. Rather, the OFT assessed whether competitors could replicate IDEXX's offer and found that they could, on the basis that many leased their in-clinic analysers, and that this could constitute a competitive alternative to IDEXX's offer.

The OFT therefore also dismissed this third area of potential concern.

Some remarks We find that the OFT's approach was based on a sound application of economic principles. As we pointed out in Section 4.5 as part of our discussion of policy implications (and more generally across all chapters of this book) the identification of a clear theory of harm is crucial, along with the key elements that would be needed to

support it and an overall compelling ‘story’ that would explain how the incumbent would have the ability and the incentive to foreclose, how the foreclosure would happen in practice, and how the incumbent would benefit from it at a later stage. In this regard, the OFT’s ‘theories of harm’ were somewhat underdeveloped and perhaps constituted more areas for investigation (or potential concern). Admittedly though, as the allegations of anti-competitive conduct were ultimately rejected, it is not surprising that the theories of harm were not particularly developed in the decision.

In our view, a few further specific aspects of the OFT’s approach are worth highlighting.

First, in carrying out the price-cost test (in a similar fashion to that discussed in Section 4.5), the OFT used an AAC benchmark, as opposed to a long-run average incremental cost (‘LRAIC’) benchmark. The OFT considered the former to be a more appropriate measure, as it was considering whether the revenue from specific contracts (accounting for a small proportion of IDEXX’s total business) was sufficient to cover the costs incurred by the supply of those contracts alone. By contrast, according to the OFT, ‘LRAIC may be more appropriate when considering rebates which cover a whole line of business, since it would then be appropriate to consider whether the revenue is sufficient to cover the attributable fixed costs associated with that line of business, and potentially some costs shared with other products, as well as the avoidable costs’ (footnote 107 of the decision).¹¹⁷

Second, the OFT examined the extent of the tied market that may have been foreclosed by IDEXX’s conduct. The implications of such an examination were very clear in the case of the bundling of the specialist tests with standard external lab tests: even if the OFT found that IDEXX’s effective price for the standard external lab tests was negative (and thus below cost), the fact that only a very limited proportion of the standard external lab tests market would have been foreclosed due to such bundles was sufficient evidence that the bundling could not raise material competition concerns.

Third, the OFT’s approach to the assessment of IDEXX’s conduct on after-markets (consumables) dealt with two important issues. One is that, when contracts are negotiated individually, a price-cost test may have

¹¹⁷ The OFT nevertheless noted that in practice the two cost benchmarks may have been similar in this case, as it expected that over a period of three years or more, it was unlikely that a significant proportion of fixed or common costs would be avoided absent those contracts.

to be applied at a contract level and not on an average basis (see also our discussion of this issue in Chapter 2, in relation to rebates more generally). The other is that the nature of after-markets may mean that it is appropriate to define a 'systems market', particularly if there is evidence that customers by and large engage in whole-life costing, and then carry out the competitive assessment accordingly.

Vertical Foreclosure

5.1 Introduction

Setting the stage Most industries are characterised by the existence of several production and distribution levels. Inputs are combined and turned into outputs. An input can be a commodity (for example, coal, iron ore, cacao) or some infrastructure (for example, an aircraft stand at an airport, a railway station, a port). An input does not have to be a physical product or facility: it could be a patent, a license, some computer code or a database, for example. Further, in retailing, an input may even amount to an end-product: retailers buy a garment, an oven, a pack of cereals – which are inputs to them – before selling them on to end-consumers.

Although a product may involve several supply levels before reaching a final consumer (for example, sourcing raw materials, manufacturing, assembling, wholesale distribution, retail distribution), for the rest of the chapter, for simplicity, we will just assume that these vertical chains comprise two levels: one upstream and one downstream.

Further, in this chapter, we will focus on industries where there is just one (or very few) upstream suppliers. Economists sometimes refer to this situation as an ‘upstream bottleneck’. Natural monopolies may fall within this category.

Moreover, we will assume that the upstream monopolist is also active downstream, with its own affiliate (for example, the owner of a port also provides its own ferry services). Unless otherwise stated, we will refer to the incumbent as I , to its upstream affiliate as U_I and to its downstream affiliate as D_I . We assume that the incumbent faces a downstream rival, D_R , which may already be in the market, or is considering entering. In the models we present in this chapter, we will mainly (with some exceptions) assume that

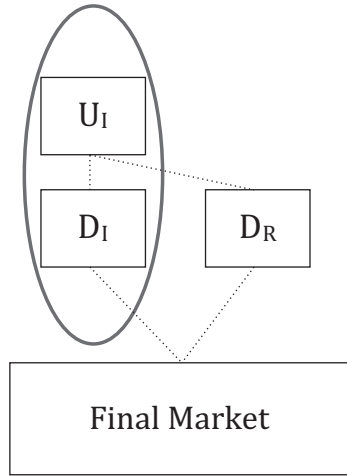


Figure 5.1. Market structure.

this rival is as efficient as or more efficient than the incumbent. Figure 5.1 depicts this highly stylised industry.

This chapter will be mostly devoted to the discussion of how economic theory may explain possible conditions under which the incumbent would rather exclude a downstream competitor, leading to vertical foreclosure (or more simply to ‘foreclosure’). As we shall see, in most cases the incumbent would have the *ability* to exclude downstream rivals, but not the economic *incentive* to do so.

When we refer to vertical foreclosure, this does not need to be complete. Even *partial* foreclosure – defined as a situation where independent downstream firms pay a higher price for the input as compared to downstream affiliates of the vertically integrated firm, but still buy a positive amount of it or still receive the input but at a degraded quality – may harm consumer welfare (see Section 5.2.3.4). This would occur if consumers ended up paying higher prices, or facing lower quality products, than in a scenario without partial foreclosure.

Finally, we note that the bottleneck does not have to be upstream for foreclosure to be possible: the bottleneck may actually be downstream (with the upstream segment being more competitive). For instance, there could be one single buyer facing several potential suppliers (think of a monopoly grocery retail chain in a small country buying from many small growers and farmers, or a Ministry of Defence purchasing military

equipment and vehicles). To a large extent, in this chapter, we shall focus on upstream bottlenecks, and we shall try to indicate where the economic implications would be somewhat different in case of downstream bottlenecks.¹

Some words of caution Following this brief discussion, an immediate reaction may be: what makes the incumbent a monopolist over the upstream input? The answer to this question will depend on the nature of the industry, as well as on the way one may define the relevant market: for example, to what extent can that input be substituted by another one? At what cost? Over what timeframe? Certain inputs or infrastructure may be unique and may not be duplicated (unless at prohibitive cost): that is, in the short- to medium-term, it may be that there is no viable alternative, no scope for multiple suppliers or providers. In some industries, matters are complicated by the presence of competition *for* the market, or winner-takes-all, for example due to the innovative nature of the products.

And even if an incumbent is truly a monopolist over the upstream input, a number of follow-on questions arise: do downstream rivals have a ‘right’ to be granted access to that input? If so, on what terms? Would final consumers be better off following external intervention (for example, by a regulator or a competition authority) that tinkers with the incumbent’s conduct? What would be the effect on the incentives of the incumbent (and possibly of firms in other industries) to keep investing and/or innovating, following an intervention which obliges the incumbent to give access to an input which may well be the product of its investments or business acumen?

We will come back to some of these questions over the course of this chapter.

At this stage, we simply note that, in our view, firms should be generally free to deal with the partners they wish. Indeed, this is also the position that the European Commission took in its Guidance Paper on how it

¹ In our dynamic vertical foreclosure model of Section 5.2.4, for instance, the implications are identical whether there is a monopoly upstream or downstream. But in other models implications may differ. Rey and Tirole (2007) – referring to a model of ‘opportunistic behaviour’ like the one we analyse in Section 5.2.3.2 – point out how upstream bottlenecks should be preferable from the consumers’ perspective to downstream bottlenecks, in that the former likely lead to lower final prices. On the other hand, downstream buyer power may be thought of as generally having a beneficial effect: buyer power would tend to decrease input prices, which in turn should – at least to some degree – be passed on as lower downstream prices. For a survey on the effects of buyer power, see Inderst and Mazzarotto (2007).

would enforce abuse of dominance provisions (in relation to exclusionary practices);² as well as that of the US Supreme Court in *Trinko*.³

It is only under very specific circumstances that there may be exceptions to this general principle, and this chapter seeks to identify such circumstances.

Practices that may fall within the vertical foreclosure framework of this chapter Vertical foreclosure may take a number of different forms in practice, which might or might not be equivalent with respect to their effects. The most obvious practice that falls within the framework of this chapter is a refusal to deal (or to supply, or to license in case the input is protected by intellectual property rights). The monopoly owner of a necessary input may simply refuse to sell it to downstream competitors and keep it all for its own downstream affiliate.⁴

Equivalently, it could set a prohibitively high price for it (a so-called constructive refusal to supply): the economic effect would be the same as an outright refusal to supply, as the downstream entrant or competitor would not find it profitable to accept such an offer.⁵

Alternatively, the incumbent could set a combination of ‘high’ upstream (or wholesale) prices and ‘low’ downstream (or retail) prices such that an (efficient) downstream competitor could not profitably operate in the downstream market. This is the notion of margin (or price) squeeze. We will discuss a few such cases in this chapter. In Section 5.2.3.1, we will also discuss why margin squeeze and refusal to supply are not necessarily equivalent practices.

² European Commission (2009). See in particular para. 75.

³ ‘Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities’ (see part III of the Supreme Court’s Opinion in *Trinko*).

⁴ Refusals to supply can also be adopted by a firm as a possible way to impede arbitrage (or parallel imports) among buyers (or countries), and thus enforce price discrimination. But this behaviour would not be ‘exclusionary’ within the meaning of this book, so we shall not dwell on it. A leading European case is *Syfait*. For further discussion, see, for example, Whish and Bailey (2015: 752–3).

⁵ In some jurisdictions, including the EU, South Africa and some of the States in the US, this practice could also fall foul of excessive or unfair pricing provisions. These laws deal with exploitative behaviour by firms with market power, which is not part of the subject of this book. However, to the extent that this amounts to the excessive pricing of an input and the (dominant) supplier is also active downstream, then such a practice would be exclusionary, and hence fall within the scope of this chapter.

Another practice that may fall within a vertical foreclosure framework is long-term capacity booking (we will review some cases from the energy sector in Section 5.6.6). An integrated gas or electricity supplier, instead of denying capacity to a downstream competitor directly, may do so indirectly by signing long-term capacity bookings with its downstream affiliate, which may leave smaller capacities available to competitors than under a competitive outcome.

There are other possible types of behaviour which may lead to vertical foreclosure, depending on the industry at issue. We already noted at the outset how an input need not have a physical nature. Intellectual property is often a key input to downstream competitors: thus the refusal to license intellectual property rights (or to provide information that is necessary for interoperability), or to do so in return for very high royalty fees, could affect downstream competition. For instance denial or degradation of interoperability may achieve foreclosure of downstream rivals (see Section 5.6.1 for a discussion of the *Microsoft* interoperability case).⁶ In yet other cases, a possible abuse may consist of a firm that owns an essential patent in a technological standard refusing to grant licences to downstream rivals, charging royalty rates that are ‘unfair’, ‘unreasonable’ or ‘discriminatory’, or using the judicial system in an anti-competitive fashion, for example through an injunction, under certain conditions (in this respect, we briefly review the *Samsung* and *Motorola Mobility* cases in Section 5.5.2).⁷ Recently, competition authorities (at least in Europe)

⁶ In some cases, refusal to supply or to license, and denial of interoperability information might be used by the manufacturer of an equipment (or primary product) as a way to prevent other firms from selling spare parts, or maintenance and repair services (the secondary, or after-market product). Since these after-market cases are typically associated with tied sales, we discuss them in Chapter 4.

⁷ There also have been some antitrust cases on patent ambushes (for example, *Rambus* and *Qualcomm*), which have received a lot of attention in both the US and the EU. Companies often need to agree on a common technology in order for consumers to be able to communicate data with each other or more generally for users to benefit from network externalities (through interoperability). Companies agree on such common standards within standard-setting organisations. A patent ambush may occur when a firm does not disclose the relevant patents it has during such meetings. Once a given standard is chosen, the owner of (essential) patents would announce it has such patents and would start charging ‘very high’ royalty fees to manufacturers wishing to produce the technology with the agreed specifications. But by that point it would be too late (costly) to devise an alternative standard. However, as in these cases the upstream firms concerned were *not* active downstream (that is, in the product markets where those technologies would be used), the economics would be quite different. These cases therefore fall outside the paradigm we analyse in this chapter. Put simply, here we are interested in exclusionary behaviour; those cases were instead about potentially exploitative behaviour.

have considered in some detail whether the holding and certain uses of some data by vertically integrated firms with market power may carry the potential to generate anti-competitive effects.⁸

Roadmap for the remainder of the chapter This chapter proceeds as follows. In Section 5.2 we provide an informal review of some of the economic theories that have been proposed to analyse vertical foreclosure. In Section 5.3 we present such theories formally. In Section 5.4 we draw some possible policy implications based on the economic theories reviewed in this chapter. In doing so, we also consider some of the possible objective business justifications and efficiencies that may exist in the context of vertical foreclosure cases. In Section 5.5, we provide an overview of some of the key US and EU cases. Finally, in Section 5.6, we discuss a few selected cases more deeply, offering an interpretation of these under some of the economic theories we set out in this chapter.

5.2 The Economics of Vertical Foreclosure

In this section, we provide a brief informal discussion of some economic theories which have focused on vertical foreclosure. We first discuss some of the arguments that have been traditionally used to justify intervention when an upstream firm owns an essential facility (Section 5.2.1). We then set out the Chicago School critique, which pointed out that the owner of an essential facility or of an essential input has the *ability* to exclude downstream rivals, but rarely will it have the *incentive* to do so in the

Moreover, these cases arguably fall more within the scope of contract law than within that of competition law. For further information on the European cases, see European Commission, Press Release IP/09/1897, 9 December 2009, for *Rambus*; and European Commission, MEMO/09/516, 24 November 2009, for *Qualcomm*. For the US, the reader is directed to *Rambus v Federal Trade Commission* and *Broadcom v Qualcomm*.

⁸ For example, EU Competition Commissioner Vestager stated: '[...] privacy isn't the only issue with data. If data can help you compete, by improving your services and cutting costs, then having the right set of data could make it almost impossible for anyone else to keep up. So we need to be sure that companies which control that sort of data don't use it to stop others from competing. That's a question that we take very seriously. But it turns out that even if you hold a lot of data, that doesn't necessarily mean you have the power to stop others from competing. What matters isn't just the amount of data. It's whether you can really use it to drive your rivals out of the market' (EU Competition Commissioner Vestager, speech delivered on 9 September 2016, available on the European Commission's website). National competition authorities have also devoted some thought to such issues – see for example the work of the UK competition authority (Competition and Markets Authority, 2015) and a joint report by the French and German competition authorities (Autorité de la Concurrence and Bundeskartellamt, 2016).

presence of an efficient downstream rival (Section 5.2.2). Next, we briefly review some theories that show that an incentive to exclude may exist where – due to particular circumstances – a vertically integrated incumbent is not able to realise all the benefits from downstream competitors (so-called imperfect rent extraction – see Section 5.2.3). Finally, we show that vertical foreclosure may occur because of a mechanism based on the existence of scale economies (Section 5.2.4). This mechanism is very similar to the one emphasised in previous chapters, and it has a dynamic component: an efficient downstream rival may be foreclosed because a lack of suitable access to inputs deprives it of the scale or customer base it needs to be viable in adjacent markets or in future periods, which may lead to consumer harm.

5.2.1 Ownership of an ‘Essential Facility’

Some readers may wonder whether industries that are characterised by the stylised scenario set out in Section 5.1 (that is, with a vertically integrated incumbent and a downstream entrant or competitor seeking access to the upstream or wholesale input) should be under the oversight of sector-specific regulatory bodies. Yet, antitrust enforcers have been dealing with similar cases for a long time. In particular, the ‘essential facilities’ doctrine is thought to have been born in the US about one century ago.⁹

The very definition of an essential facility is fraught with difficulty and most attempts end up in tautological or circular definitions.¹⁰ Put loosely, the essential facilities doctrine concerns industries where a firm controls some upstream infrastructure or input without which a competitor cannot operate downstream.¹¹ Typically, for the essential facilities doctrine to potentially apply, the owner of the facility is also active downstream, where the potential competitor wants to enter, and the costs of duplicating the upstream facility are too high.¹²

⁹ Admittedly, though, the US Supreme Court has never endorsed it or rejected it, while lower courts applied it more or less consistently until the 1980s (we review the landmark case-law in Section 5.5). The interested reader is directed to legal scholars for a more competent legal assessment of these cases, for example Areeda (1990), Jones and Sufirin (2016) and Whish and Bailey (2015).

¹⁰ For further reading, see Motta (2004: section 2.5.2), O’Donoghue and Padilla (2013: section 10.1), Ridyard (1996) and Werden (1987).

¹¹ Or, conversely, the firm owns a downstream facility without which competitors cannot compete upstream.

¹² If the facility-owner was only active upstream, a refusal to deal coupled with exclusivity with one downstream firm only would be less suspicious in terms of exclusionary intent (since the facility-owner is not active downstream); yet it could be a way of restoring

While this setting looks fairly straightforward to describe, the difficulties with this ‘doctrine’ arise in the details of its application and in the determination of what conditions need to be satisfied for authorities to intervene. In applying the essential facilities doctrine in *MCI*, determining that it was illegal for a local telecommunications monopolist to deny interconnection to its own local network to competing long-distance carriers, a US Appeals Court clarified the set of cumulative conditions that are necessary for a refusal to deal to infringe US antitrust law:

(1) control of the essential facility by a monopolist; (2) a competitor’s inability practically or reasonably to duplicate the essential facility; (3) the denial of the use of the facility to a competitor; and (4) the feasibility of providing the facility. (Para. 192 of the judgment)

Werden (1987) suggested that the above test raised a number of questions: what constitutes a ‘facility’? When is it ‘essential’? What exactly is a ‘monopolist’ (that is, does the test not catch a firm with substantial market power)? What does ‘control’ exactly mean (for example, is the exercise of contractual rights caught, as opposed to simple outright ownership)? What constitutes a denial of use of access (for example, does it include offering access on ‘reasonable terms’ and if so what would this mean)? At what point does access become ‘not feasible’? Why are potential defences (for example, legitimate business justifications) missing from the above test? Why are *ex ante* incentives to invest not taken into account?

Werden (1987) added an interesting remark, from a broader policy perspective, with which we agree. Notwithstanding the theoretical merits (or lack thereof) of the essential facilities doctrine, one cannot apply it to a given case unless there are some workable, fair and proportionate remedies. Moreover, it should be applied only where the market conditions are sufficiently ‘extreme’ for this to make economic sense, such as in the case of (most) natural monopolies. In essence, there appears to be an overlap between the market conditions that may call for the application of this doctrine and those typically associated with a need for sectoral regulation. In other words, either a given industry needs a regulator (which will ideally know the industry concerned in great detail and be able to set appropriate access terms if needed) and then the essential facilities doctrine

monopoly profits – see Section 5.2.3.2 for a discussion of the so-called ‘commitment problem’.

is somewhat redundant; or competition is broadly viable in an industry, and the essential facilities doctrine does not apply.¹³

As we will see in Section 5.5 where we review the landmark case-law on refusal to deal, the essential facilities doctrine did not survive much beyond the 1980s in the US, while it in fact began to be applied at EU level at the beginning of the 1990s.

The relative decline of the essential facilities doctrine in the US corresponded to the period when the lessons of an influential school of thought were beginning to influence antitrust enforcement: it was the Chicago School critique, to which we turn next.

5.2.2 The Chicago School Critique

Consider the stylised industry setting we presented in Section 5.1, in Figure 5.1. At an intuitive level, one could conjecture that the incumbent would like to extend its upstream monopoly downstream and earn higher profits. Can it do so? Legal considerations aside, many of the practices we listed in Section 5.1 would actually enable the incumbent to foreclose the (existing or potential) downstream competitor. So, trivially, foreclosure is almost always feasible in such a setting. But the key question is: would it be *profitable* for the incumbent to foreclose the downstream competitor? The short answer given by Chicago scholars is: ‘no’.

This logic is very similar to the argument put forward by the Chicago School in the case of tying (reviewed in Chapter 4). Assume the downstream competitor (or entrant) is at least as efficient as the incumbent, either because it can offer the good at equal or lower cost or because its version of the good generates at least as much value to consumers.¹⁴ For simplicity, also assume that there are no capacity constraints and that there is a one-to-one transformation technology: one unit of input is turned into one unit of output downstream (by either the incumbent or the competitor). Then the incumbent can offer the downstream competitor a contract which has the following features: it will be accepted by the downstream rival (that is, the terms are such that staying out of the market

¹³ Perhaps one area where this appealing syllogism does not carry through is with intellectual property. There are no sectoral regulators in each industry that has intellectual property. But this should not imply that any refusal to supply intellectual property should be legal (see Shelanski, 2009 for a deeper discussion). More generally, there may also be situations where there should be a regulator, but for some reason there is none.

¹⁴ The assumption that the competitor is as efficient as the incumbent allows to focus on situations where vertical foreclosure – if it takes place – may be detrimental for welfare.

would not be better); and it will allow the incumbent to extract the rival's 'efficiency rent' (that is, put very loosely, allow the incumbent to turn the downstream rival's cost advantage into revenues for the incumbent). In these circumstances, the incumbent is better off serving the competitor rather than excluding it.¹⁵

A simple example Consider the following (very stylised) example, referring back to Figure 5.1. Incumbent I has an upstream affiliate U_I which produces an input at zero cost and a downstream affiliate D_I that transforms it into an output at unit cost c . There is a downstream rival, D_R , which transforms the input at zero cost into the same output. Suppose there is only one consumer (to keep the example as simple as possible) who buys at most one unit and whose maximum willingness to pay for the good is 1. Suppose the vertically integrated incumbent I refuses to supply D_R . Then I will be a downstream monopolist, and will charge the maximum price at which the consumer will buy (that is, 1), thereby earning a profit equal to $\pi^{vf} = 1 - c$.

But I would earn more by actually serving D_R and earning upstream profits from selling the input to it. To appreciate this, suppose, for instance, that I proposed to D_R a contract whereby it sells the input at a unit price w and it commits not to compete with D_R through its own subsidiary (we shall discuss this assumption in Section 5.2.3.2). Then D_R would be monopolist in the downstream market and would sell at the price of 1, by making a profit $1 - w$. Should D_R reject such contract, it would make zero profits, so it has an incentive to accept any contract that allows it to make a positive (or at least a non-negative) profit. The optimal choice for firm I would therefore be to set $w = 1$. In this way, all the profits made by D_R would be 'extracted' by the incumbent firm.¹⁶ In other words, by serving

¹⁵ This is also called the 'single monopoly profit theory', for the following reason: there is only one market whose monopoly profits may be reaped, and this is the final market; the upstream firm is already able – through an appropriate choice of the contract it offers to the independent downstream firm (or firms) – to extract all the monopoly profits from this market. Indeed, Rey and Tirole (2007) stressed how vertical foreclosure in a static framework should not be seen as a way of 'leveraging' market power from one market to another (for example, from the upstream market to the downstream market) but rather as a way of protecting or capturing that very single monopoly profit.

¹⁶ Note that D_R makes zero profits both if it accepts and if it rejects the offer. For simplicity we assume that when indifferent it chooses the former. But note that this is not a strong assumption: it would be enough to say that w is equal to 1 minus one cent, and D_R would strictly prefer accepting the offer.

the independent rival, firm I earns profit $\pi^{novf} = 1$, which is higher than the profit $\pi^{vf} = 1 - c$ it earns if it forecloses.¹⁷

Some final remarks In sum, the argument of the Chicago School is that, while foreclosure is always (trivially) feasible for an incumbent which is an upstream monopolist, it is not profitable. So if we observe foreclosure in reality, this is likely to arise because the competitor is less efficient, or because there are substantial efficiencies for the incumbent to organise production and distribution internally (that is, to use its own upstream and downstream affiliates together).

By emphasising the fact that an incumbent must have an incentive – and not only the ability – to exclude, the Chicago School makes a valid and important point. When vertical foreclosure is at issue, one always has to investigate whether there is an incentive to exclude and whether exclusion would be welfare-detrimental, that is, whether there is a valid theory of harm.

However, there exist a number of circumstances that limit the incumbent's ability to extract rents from the downstream market and these *may* give rise to the incentive to foreclose the rival (Section 5.2.3). Moreover, the standard framework considered by the Chicago School is static (that is, it focuses on a single period). Section 5.2.4 will show that incentives to foreclose may also arise in the presence of a dynamic framework, in which current decisions may affect future market outcomes.

5.2.3 Imperfect Rent Extraction

In the simple example in Section 5.2.2 where we set out a stylised version of the Chicago School critique, we assumed that the vertically integrated incumbent could set whatever price it wanted for its input and this would allow it to extract all rents from the final market.¹⁸ However, there may be circumstances under which the incumbent will not be able to extract all the

¹⁷ Note that foreclosure would never arise even if it was D_R which made the contractual offer to firm I . This is because firm I has always the option to refuse the contract and earn profits $\pi^{vf} = 1 - c$. Hence, firm D_R will have to choose a w such that the incumbent will make at least as much as π^{vf} . Generally, for intermediate degrees of bargaining power, firm I will be strictly better off serving than excluding firm D_R . Note that unless the bargaining power is all on the incumbent, there will be imperfect rent extraction. However, as this example shows, imperfect rent extraction is not a sufficient condition for the incentive to foreclose to arise.

¹⁸ Typically, the incumbent would choose a non-linear price, composed of a fixed fee plus a variable component, although in the example we presented with rigid demand the latter component was sufficient for full rent extraction by the vertically integrated incumbent.

rents from the downstream market: this may provide it with an incentive to foreclose downstream rivals.

In what follows, we shall describe some of such circumstances at an intuitive level. (See Section 5.3 for a more technical treatment.) Section 5.2.3.1 analyses the case where a sectoral regulator restricts the incumbent's freedom to contract with the downstream rivals; Section 5.2.3.2 shows how opportunistic behaviour may prevent the incumbent from extracting monopoly rents from the market; Section 5.2.3.3 investigates the effects that may arise in a context with uncertainty and risk-averse downstream firms; finally, Section 5.2.3.4 looks at the case where the incumbent faces some competition in the provision of the input and engages in refusal to supply the input so as to increase the costs of the downstream rival.

All the cases discussed in this section have imperfect rent extraction as their common trait. Note however that imperfect rent extraction is not by itself sufficient for vertical foreclosure to take place. For an incentive to exclude to exist, it must also be that by only relying on its downstream affiliate, the incumbent would earn higher profits than if downstream rivals were active. This is not always the case.¹⁹

5.2.3.1 Regulation

One of the reasons why a dominant supplier of an input may be at least partially unable to extract profits from a market is that it may be subject to some form of regulation.²⁰ Certain firms are vertically integrated and operate in industries where there is typically a sector-specific regulator. Think, for instance, of a railway network, of a fixed telephony network, of distribution and transmission in energy and so on. In such cases, a firm may be a monopolist (or have significant market power) upstream while at the same time have an affiliate which is one of the firms active in

¹⁹ For instance, earlier in this chapter, in footnote 17 above, we showed that if the incumbent does not have full bargaining power, there will be imperfect rent extraction (some of the downstream profits have to be left to the downstream rivals), and yet supplying efficient downstream rivals generates larger profits for the incumbent than if it foreclosed them. Similarly, as we shall see below, the very fact that there is a sectoral regulator limiting access prices is not sufficient for vertical foreclosure. Furthermore, under elastic demands and linear pricing, supplying a more efficient downstream rival gives more profits than own production, even if it is not possible to extract all the rival's rents. Finally, the very presence of upstream rivals who offer the (possibly inferior) input does not necessarily create an incentive for foreclosure: the presence of other input suppliers also means that downstream rivals can get the input from other sources, so that profits under foreclosure might be lower than otherwise.

²⁰ In Chapter 4, we provide an analogous discussion in relation to tying.

the (fully or partially unregulated) downstream market: for instance – to continue with the example above – by operating passenger train services, offering internet and telephone services, or being a retail supplier of energy to households and businesses.

The sectoral regulator will typically want to ensure efficient pricing of the upstream input (while keeping the incentives to invest upstream), and to promote fair competition downstream. This may limit the ability of the upstream monopolist to extract profits from the sale of the input, and may potentially induce it to try to foreclose downstream rivals and favour the own downstream affiliate instead.

In the technical Section 5.3.2 we show a simple illustrative example in which the regulator imposes a cap on the wholesale unit price that the upstream affiliate of the incumbent can charge to the downstream firms. This model shows that if the cap is sufficiently low, the rents extracted from the more efficient downstream rival are so low as to make the vertically integrated incumbent prefer (legal and regulatory considerations aside) to engage in vertical foreclosure by refusing to supply the input and to monopolise the downstream market through the less efficient own affiliate. Importantly, this model also shows that the less efficient the downstream affiliate, the lower the profits obtained when the vertically integrated firm monopolises the downstream market, and the less likely is foreclosure. In the model we present in this chapter, we focus on differences in production costs between the own affiliate and the downstream rival.²¹ A similar argument would apply in the presence of horizontal product differentiation (that is, when consumers differ in their taste or in their geographical distance from the sellers): the more pronounced the differentiation between the product sold by the incumbent's downstream affiliate and the one offered by the downstream rival, the less profitable to monopolise the final market through the own affiliate (because a great proportion of demand would be lost), and the lower the incentive to engage in vertical foreclosure. The same logic applies under vertical product differentiation (that is, when firms differ in the quality of the products they sell): the higher the quality of the rival relative

²¹ The model we shall present abstracts from a number of potential complications that may be relevant in certain industries. For example, while we assume that the upstream input can be used seamlessly across the downstream affiliate of the vertically integrated incumbent and any independent downstream competitors, in practice, in some circumstances, the vertically integrated incumbent may incur additional upstream costs in order to supply downstream competitors. For a model that considers this, see, for example, Bouckaert and Verboven (2004).

to the downstream subsidiary, the less profitable to resort to refusal to supply, and hence the less likely a foreclosure outcome.

Degraded access Of course, regulatory obligations (and competition law provisions) may make it difficult for the incumbent to exclude downstream rivals by engaging in refusal to supply: the sectoral regulator may not allow it to refuse the input to downstream rivals, for instance. However, even if outright refusal to supply was not possible, the upstream monopolist may resort to more subtle ways to foreclose the downstream rivals. For instance, it may reduce (an unverifiable component of) the quality of the input, or degrade interconnection, or claim that it does not have enough capacity to serve them, or delay the supply of the input. We believe it is not a coincidence that many of the cases of vertical foreclosure which have been analysed by competition authorities and courts took place in sectors whose upstream segments were regulated, and often took the form of more or less circuitous ways to reduce availability and quantity of input supply or increase the cost of access for downstream rivals (see the review of the case-law and the discussion of selected cases in Sections 5.5 and 5.6).

Margin squeeze In some cases the incumbent may engage in a *margin squeeze*. In technical Sections 5.3.2.1 and 5.3.2.2 we will show that refusal to supply and margin squeeze should not be regarded as equivalent practices in a regulated environment.

By refusing to sell the input (or by setting a prohibitively high price for it) the vertically integrated incumbent will make it impossible for the rival to operate in the downstream market and, at the same time, it will manage to set the monopoly price in that market. If the vertically integrated firm cannot engage in refusal to supply – for instance because the presence of regulation forces the upstream monopolist to sell the input at a regulated price – then it may engage in margin squeeze, that is, it may choose a combination of the wholesale price and of the retail price such that the independent rival cannot profitably operate in the downstream market. However, to do so, the retail price must be distorted downwards, thereby reducing the profits that the incumbent obtains by engaging in vertical foreclosure and making it less likely.²²

²² The evolution of the EU competition law is *not* in line with this insight. The Court of Justice's judgment in *TeliaSonera* (which we describe in more detail in Section 5.5.2) has established that margin squeeze may represent an abuse even in the absence of input indispensability, whereas input indispensability is a necessary condition for refusal to supply. Hence, the practice which is more likely to raise a concern, namely refusal to supply, requires a stronger condition for a finding of an infringement.

Importantly, the distortion of the retail price that forecloses the activity of the independent rival involves a profit sacrifice for the vertically integrated incumbent: in the short-run it would find it more profitable to supply the downstream rival rather than to engage in margin squeeze.²³ Hence, for margin squeeze to be an optimal strategy the profits lost in early periods (relative to the case in which the independent firm is supplied) need to be recouped in later periods. Recoupment will be possible if the vertically integrated incumbent will see its market power increase in later periods. Then, it must be that a squeeze of the rival's margin in early periods leads to exclusion (or marginalisation) of the downstream rival in later periods, for instance because the downstream market features scale economies that the incumbent's affiliate has already exploited while the downstream rival has not (like in the model of predation developed in Chapter 1). Section 5.3.2.2 will analyse the incentives to engage in margin squeeze in this type of environment. Alternatively, there must be financial frictions and asymmetries between the incumbent's affiliate and the downstream rival in terms of ability to raise external funds. From this perspective, predation and margin squeeze show many similarities.²⁴ We will come back to this issue in Section 5.4.2, where we discuss policy implications.

5.2.3.2 Opportunistic behaviour

When we presented the Chicago School critique in Section 5.2.2, we implicitly assumed that – at the moment of contracting with the downstream rival – the incumbent is able to commit not to use its downstream affiliate to compete with the rival. However, if it was not able to do so, the incumbent would be able to extract lower profits from supplying downstream rivals than in the presence of a credible commitment. This

²³ In a setting in which downstream firms sell homogeneous products and one unit of input is transformed into one unit of output, a profit sacrifice in the short-run is detected when the margin obtained by the vertically integrated firm on the units sold in the final market falls short of the margin lost on the units of input that the vertically integrated firm fails to sell to the downstream rival when its activity is foreclosed: $p - c_{DI} - c_{UI} < w - c_{UI}$, or equivalently when $p - c_{DI} < w$. This inequality can also be interpreted in the light of replicability: $p - w < c_{DI}$ implies that a downstream rival as efficient as the incumbent's affiliate would not manage to cover downstream costs if it matched the incumbent's retail price and paid the price w for the input. On these issues see also Section 5.4.3.

²⁴ This insight is not in line either with the recent evolution of the EU competition law. In Section 5.5.2, we will see how, since at least the *Deutsche Telekom* judgment, margin squeeze has been held a stand-alone type of anti-competitive behaviour (see also footnote 22). Instead, in the US the 'low prices' leg of a margin squeeze allegation would warrant a predation-style analysis.

may in turn provide an incentive to exclude the downstream rival. This section explains the intuition behind this claim.

As the discussion in Section 5.2.2 suggested, the incumbent would like to extract as much economic rent as possible from the downstream rival through a suitable choice of contracts. This means that the incumbent would want to proceed as follows: (i) let the (more efficient) downstream rival sell all (or most) of the output at the (industry) profit-maximising price (that is, the incumbent would keep the downstream affiliate inactive); and (ii) ‘extract’ the profits thus earned by the downstream rival through a corresponding lump-sum payment to the incumbent (for example, in exchange for the input supplied).

However, if the downstream rival feared that the incumbent would in parallel use its downstream affiliate to compete for consumers (that is, to behave *opportunistically*, reneging on its promise not to be active itself in the downstream market), its willingness to pay for the input would decrease, since expected competition from the incumbent’s downstream affiliate would decrease the downstream rival’s expected profits. In turn, this would limit the ability of the incumbent to extract profits from the downstream rival through the lump-sum payment described and this may make a foreclosure strategy potentially more profitable than the one set out above.²⁵

Note that the inability to commit not to operate the downstream affiliate does not necessarily lead to *complete* foreclosure of the downstream rival. For instance, Reisinger and Tarantino (2015) show that, if the downstream rival is more efficient than the incumbent’s downstream affiliate and the efficiency gap between the two firms is not too large, then the incumbent would engage in *partial foreclosure*: it supplies the independent rival but on less favourable terms (that is, at a higher wholesale price) than the own affiliate.²⁶ By so doing the incumbent earns lower profits than in the scenario with complete rent extraction, but such profits are larger than the ones obtained by excluding the more efficient rival entirely.

Opportunistic behaviour in Hart and Tirole (1990) The fact that the monopoly supplier of an input may suffer from opportunistic behaviour limiting its ability to extract rents from a downstream market has been first proposed by Hart and Tirole (1990). They refer to a different situation,

²⁵ In Section 5.3.3, we will present a very simple model to illustrate the impact of the inability to commit not to operate the downstream affiliate.

²⁶ Instead, if the efficiency gap is large enough, the incumbent finds it optimal to offer to the downstream rival a wholesale price that is indeed lower than the one paid by the own downstream affiliate.

though, in which the incumbent is a monopolist upstream, but it is not vertically integrated and then does not directly operate in the downstream market. In this setting imperfect rent extraction arises for a different reason, namely from the upstream monopolist's inability to commit to specific contractual terms.²⁷

In Chapter 2 (specifically, in Section 2.3.4.1) we illustrate their argument through a stylised example, which we repeat here for convenience. Assume that there are two independent downstream firms, D_1 and D_2 , which are equally efficient and sell a (imperfectly) substitutable final product. Crucially, assume also that the terms of the contract that the upstream monopolist offers to each downstream firm cannot be observed by the other downstream rival.²⁸ As shown below, the upstream monopolist has an incentive to agree on better terms with one downstream rival to the detriment of the other, and this limits its ability to extract monopolistic profits from the downstream market (since the two downstream firms are aware of the incumbent's incentives).

Suppose that the retail prices p_1^* and p_2^* with the associated sales q_1^* and q_2^* are the ones that a firm vertically integrated with both downstream units would choose so as to maximise industry profits. Imagine also that there exists a wholesale price w^* such that, if D_1 and D_2 paid w^* for each unit of the input, then they would set p_1^* and p_2^* in the final market, selling q_1^* and q_2^* and making profits equal to $\pi_1^* = (p_1^* - w^*)q_1^*$ and $\pi_2^* = (p_2^* - w^*)q_2^*$. The upstream monopolist U_I could then extract all the maximal industry profits if it was able to convince D_1 and D_2 to accept a contract whereby they buy each unit at a price w^* and in addition pay a fixed fee π_i^* to U_I .²⁹ But the downstream firms will not want to sign such a contract. To understand why, suppose that they did sign such a contract and consider what would happen next.

²⁷ See also the subsequent work by O'Brien and Shaffer (1992), McAfee and Schwartz (1994) and Rey and Vergé (2004). See also Rey and Tirole (2007) for an insightful review of this literature.

²⁸ Equivalently, one may assume that if such a contract was publicly observable, it could also be privately renegotiated; that is, it is impossible for the upstream monopolist to credibly and publicly commit to a certain price for the sale of the input.

²⁹ The fact that with this contract they would make zero profits is *not* what will make them reject the contract. We are assuming that the upstream firm has all the bargaining power and that if they rejected the offer, the downstream firms would not have the input, and would therefore make zero profits. So if they did make zero profits, they would still accept the contract. (If the reader is uncomfortable with the idea that the gain is exactly zero, one can posit that the incumbent asks for a fixed payment which is slightly less than π_i^* , so that D_1 and D_2 would make strictly positive profits.) The problem comes from the fact that if they accepted such contracts they would make negative profits, as explained below.

After agreeing on the above terms, with each downstream firm having paid π_i^* to U_I , the upstream monopolist would have an incentive to renegotiate and offer either firm, say D_1 , the input at a slightly lower unit price than w^* . This would allow D_1 to have a lower input cost than its rival, sell a quantity $q' > q_1^*$ (if competition was very fierce, it might even serve the whole market), and earn profit $\pi' > \pi_1^*$. Therefore, D_1 would be willing to pay as much as π' for the new (renegotiated) contract, giving the upstream monopolist an additional profit $(\pi' - \pi_1^*)$.

Note, however, that firm D_2 's profits would fall as a result of this, since the original contract commits it to pay π_2^* , but after the renegotiation between U_I and D_1 , it would sell less and earn less than π_2^* (if competition was very fierce, it would sell and earn nothing). Since D_2 will anticipate the upstream monopolist's temptation to renegotiate the contract with D_1 , it would then be unwilling to sign a contract with the upstream monopolist under which it pays π_2^* . Note that the same might happen with either of the two downstream sellers, so neither would be willing to enter into an agreement with U_I unless a very low fixed payment is set.

The standard result from the literature on commitment is therefore that a monopolist is unable to exploit all of its monopoly power, being harmed by its own temptation to change the terms with the downstream firms (that is, to behave opportunistically). Only if it was able to credibly commit to certain contract terms, would it manage to extract all monopoly profits from the downstream firms.

How can the upstream monopolist (that is, the monopoly provider of the input) solve its commitment problem? In Chapter 2 (in particular, in Section 2.3.4.1), we discussed the possibility for the upstream monopolist to commit to trade in exclusivity with one of the downstream sellers. If such a contract was enforceable in a court of law, this would guarantee that the upstream monopolist could not offer the input to any other competitor, and it would allow the incumbent to extract all the profits from the (sole) downstream seller.³⁰ Another possibility is to vertically integrate with one of the downstream sellers and then to refuse to supply the other(s). Note that if the downstream firms transform the input into homogeneous products, vertical integration combined with refusal to

³⁰ Other possible ways to solve the problem (legal considerations aside) include resorting to resale price maintenance, to most-favoured nation clauses, or simply by reputation: if U_I and the downstream firms were going to interact repeatedly over a long horizon (as may presumably be the case in many supplier-distributor relationships), then the incumbent may be able to solve the commitment problem simply by establishing the reputation of not renegotiating contract terms.

supply allows to achieve the maximal industry profits.³¹ If the downstream firms are differentiated, then vertical integration combined with refusal to supply may emerge as a second-best solution that reduces industry profits as compared to the maximal ones, but allows larger profits than in the case of vertical separation to be obtained.

Upstream oligopoly So far we have focused on the potential anti-competitive conduct by a single vertically integrated incumbent. Reisinger and Tarantino (2015) – like Hart and Tirole (1990) before them – show that the incentive to engage in (partial) foreclosure may persist in a model where a dominant input supplier faces competition from a fringe of (less efficient) alternative input providers.

Similarly, under certain conditions, even equally efficient upstream competitors may have an incentive to engage in the collective vertical foreclosure of a more efficient downstream entrant. Calcagno and Giardino-Karlinger (2016) find that if the downstream entrant is not too much more efficient than the downstream affiliates of two vertically integrated duopolists, such duopolists may have the ability and the incentive to exclude the more efficient downstream entrant. In their model the possibility of opportunistic behaviour (reneging on their promise not to be active in the downstream market through their own downstream affiliates) affects the incentives for the duopolists to tacitly coordinate as opposed to compete, so that the vertically integrated incumbents may find it more profitable to foreclose a more efficient downstream entrant than supplying the input to it.

5.2.3.3 Uncertainty and risk aversion

Another context in which vertically integrated firms may be unable to fully extract economic rents from downstream rivals is where downstream rivals: (i) are risk-averse; and (ii) face some uncertainty in the industry. Suppose for instance that a vertically integrated firm faces downstream rivals which are in expected terms more efficient than the downstream affiliate, but whose ultimate profits are uncertain. For example, each downstream firm's costs may be private information and uncertain, so that a given downstream firm's profit would be higher the higher the realisation of the rivals' costs. If the downstream firms are risk-averse, then the upstream affiliate will not be able to set contract terms such as to extract all rents from downstream competitors, because the latter will want to be

³¹ If one downstream firm is more efficient than the other(s) the upstream monopolist would have an incentive to integrate with the more efficient firm and to refuse to deal with the other(s).

compensated from the risk of facing rivals which are low costs – and thus of having low profits.³²

In technical Section 5.3.4 we study a model based on Hansen and Motta (2012) and we show that if downstream rivals are very risk-averse, then the ability of the upstream monopolist to extract rents from them would be significantly reduced. On the other hand, the fact that the rivals are more efficient would make it more willing to serve them, giving rise to a trade-off. To the extent that downstream rivals are not much more efficient in expected terms than the incumbent's downstream affiliate, vertical foreclosure will take place under this model and the vertically integrated firm will not sell any input to downstream rivals.

5.2.3.4 Upstream competition and the raising rivals' costs argument

Most of the practices referred to in this chapter are likely to result in the competitor or entrant facing a 'higher cost' as a consequence of the incumbent's behaviour. They therefore share the flavour of a strand of economic literature on 'raising rivals' costs' ('RRC'), whose original contributions are due to Salop and Scheffman (1983 and 1987) and Krattenmaker and Salop (1986). These authors identify a number of practices which may correspond to such RRC strategies. Some, like exclusive dealing, are analysed in other chapters. Others, like engaging in a race in innovation, investments or advertising, we shall not discuss because we would not expect them to be anti-competitive in the first place; others still, like outright sabotage and vexatious litigation, would be good examples of very direct raising rivals' cost strategies, but are probably infringing other laws before competition law.^{33,34}

³² A risk-neutral agent may be willing to accept a contract whereby it has to pay a fixed fee which is equal to its expected gains, but a risk-averse one will never accept such an offer. In the limit, an infinitely risk-averse agent's expected utility corresponds to its worst possible outcome, and will not be willing to pay anything more than the expected earnings under that outcome.

³³ In some recent cases of abuse of dominant position in the pharmaceutical sector, though, there may have been infringement of competition law but not necessarily of regulatory or patent laws. In *Ratiopharm/Pfizer*, the Italian competition authority found Pfizer to have engaged in a series of practices – such as filing patents without merit and engaging in spurious litigation – aimed at delaying the entry of a generic version of Xalatan, Pfizer's originator product. Likewise, the French competition authority found Sanofi-Aventis to have denigrated the competitors' generic versions of a drug containing active ingredient clopidogrel and made false statements to discourage distributors and pharmacies from selling them. Finally, in *AstraZeneca*, the European Commission fined AstraZeneca for having made an improper use of the patent and of the regulatory systems to pre-empt generic entry.

³⁴ One example of a case investigated by the European Commission concerns railways (*Baltic rail transport*). The Commission opened proceedings against the Lithuanian railway incumbent AB Lietuvos geležinkeliai, as it suspected that this firm had limited

But several practices that a vertically integrated firm may engage into may well belong to the category of RRC strategies. For instance, as we already discussed in Section 5.1, the upstream affiliate may increase the cost of its downstream rivals by engaging in input degradation, by making access to its input otherwise difficult, by price discriminating, setting an excessive price for the input, or by making a straightforward refusal to supply. In this section, we shall analyse in particular refusal to supply (which can be seen as the extreme version of the above-mentioned practices) and show how it can be both rational and have an anti-competitive impact. Interestingly, though, it turns out that its effect *within this literature* is not exclusionary, but rather aimed at relaxing competition in the marketplace.

The main reference in this respect is Ordober, Saloner and Salop (1990), henceforth 'OSS'. They consider an industry in which a vertically integrated firm faces both an upstream competitor and a downstream competitor, and suppose that the vertically integrated firm is able to commit not to supply the downstream rival. In that case, the latter will end up paying a higher price for its input requirements, since it will have to buy the input from the independent upstream firm, which will enjoy monopolistic power (or more generally stronger market power) over the independent downstream firm when the integrated incumbent commits not to supply the input to it.

In other words, by refusing to supply the independent downstream rival, the vertically integrated firm increases the costs of the downstream rival, which will be less competitive (and hence less aggressive) in the downstream market, resulting in higher downstream prices and higher downstream profits for the vertically integrated firm. Note that the downstream competitor is not completely excluded from the market: it just faces higher input costs, but the practice is still anti-competitive because it leads to higher final prices and lower consumer welfare.

The OSS result crucially relies on the integrated firm being able to (publicly and irreversibly) commit not to serve the competing downstream firm, an assumption which may be relatively difficult to maintain. However, some subsequent work has suggested that in particular cases there may be credible ways for the vertically integrated firm not to supply independent downstream rivals, for instance by designing an input which is compatible

competition on the rail markets in Lithuania and Latvia by dismantling a railway track connecting the two countries. This may have prevented customers from using other rail operators for the transport of freight between Lithuania and Latvia. See European Commission, Press Release IP/15/2940, 5 January 2015 (as well as updates in case COMP/39.813).

with its downstream affiliate only (see Choi and Yi, 2000; Church and Gandal, 2000).

Allain et al. (2016) take the intuition of OSS one step further, and show that a vertically integrated firm may have the incentive to act in such a way as to expose the independent rivals to hold-up by alternative upstream suppliers, thereby reducing their incentives to invest, and their competitiveness. Consider the same setting as OSS, with a vertically integrated firm facing two independent rivals, one upstream and the other downstream. Suppose also that the downstream firms can invest to improve their product (or reduce their cost), but investment decisions take place before the contractual terms between the upstream and downstream firms are set. In this setting, once the payment for the input provision is made, the upstream affiliate has an incentive to degrade the quality of the input that it supplies to independent rivals – or more generally to degrade non-contractible components associated with the input provision – because this benefits its own affiliate. For instance, input provision may allow the upstream affiliate to obtain crucial information concerning the independent firms' business that can be used to the benefit of the own affiliate and to the detriment of the independent firm supplied. The expectation of input degradation by the upstream affiliate allows the alternative upstream supplier to secure input provision to the independent firm at favourable conditions, that is, by extracting large profits from it. In turn, as the independent downstream supplier expects that it would only be able to appropriate small rents in the *ex post* input negotiation, its incentive to invest is weakened, to the benefit of the downstream affiliate of the integrated company.

We provide further discussion of OSS in the context of a model presented in technical Section 5.3.5.

5.2.4 A Dynamic Theory of Vertical Foreclosure

In Chapter 4, we present a dynamic leverage theory of tying, due to Carlton and Waldman (2002) and inspired by the *Microsoft* case in the US. Under their model, an incumbent firm which is a monopolist of a primary product, and also sells a complementary one, may want to bundle the two products, thereby excluding a rival producer of the complementary product. This is not done to achieve monopolisation of the complementary market, but in order to protect the monopolistic position in the primary one where the incumbent faces future entry.

In a static context, tying would not be optimal: the incumbent would make more profits by allowing sales of the rival complementary product

and extracting profits through appropriate pricing of the primary product (because of the complementarity assumed in the model, a consumer who buys the rival complementary product always buys the incumbent's primary product). But in a dynamic context, it may be rational to forgo short-run profits in order to keep a monopolistic position in the primary market.

Given that markets that are vertically related share similar features from an economic perspective with complementary markets, Carlton's and Waldman's theory may be extended to explain vertical foreclosure as well.³⁵ In technical Section 5.3.6 we offer a simple formalisation of this insight and also show that dynamic vertical foreclosure may take other forms (for example, not just by protecting monopoly power in the primary market but also to transfer it to the complementary market). Next, we present the same ideas informally.

Dynamic foreclosure to 'protect' a monopolistic position Consider a situation in which there is a vertically integrated incumbent which has a monopoly both upstream and downstream, and suppose that it currently faces entry by a more efficient downstream rival, while in a future period it faces potential entry by a more efficient upstream rival. Assume there is a single homogeneous good. From a static perspective, that is, ignoring future entry by the upstream rival, the incumbent would typically prefer dealing with the downstream rival: since the latter is more efficient, if the incumbent's upstream affiliate trades with it, the incumbent will typically earn higher profits (consistently with the Chicago School argument reviewed in Section 5.2.2).

However, if the downstream rival entered the market at a given point in time, then the upstream rival would find it easier to enter the upstream market in the future. Under the stylised model proposed, if both the downstream and the upstream rivals are in the market, the (less efficient) incumbent will make no profits.³⁶ Hence, from a dynamic perspective, that is, to prevent future upstream entry, under this model the incumbent may want to foreclose current downstream entry in the first place. Like in Carlton and Waldman's model, this is not because of a desire to monopolise the downstream market, but it is to protect monopolistic upstream profits.

³⁵ See also Fumagalli and Motta (2017b).

³⁶ In the model discussed in Section 5.3.6, we shall assume that products are homogeneous and that firms compete in prices, so the incumbent's profits will be exactly zero. More generally, though, when both upstream and downstream entry takes place, the incumbent will make lower profits due to competition at both stages of production.

Note also that the same logic applies if the vertically integrated incumbent faced upstream entry at a given point in time and downstream entry in the future. In that case, exactly the same dynamic incentive to foreclose may arise. Indeed, the incumbent may refuse to buy from the more efficient upstream entrant at a point in time in order to prevent future downstream entry in the future. (And again, from a static perspective, foreclosure of the upstream rival would not take place, because by buying from it the incumbent would make higher profits in the short-run than by refusing to deal with it.)

Dynamic foreclosure to ‘transfer’ monopoly power Consider now the same situation as above, namely a vertically integrated incumbent facing entry by more efficient rivals both upstream and downstream. However, suppose that either will enter for sure (for instance, it might be impossible to prevent its entry because it has very low fixed setup costs relative to the profits it could make even in the worst-case scenario).

To fix ideas, suppose that the downstream entrant might enter at a given point in time; while the upstream entrant will enter at a later date for sure. In this case, the incumbent knows that it will lose the upstream monopoly in any event, but it may want to foreclose the downstream entrant at present (even if unprofitable in the short-run), so as to obtain a downstream monopoly position in the future. If it dealt with the downstream entrant at present, the incumbent would face rivals both upstream and downstream in the future, thereby obtaining zero profits under the conditions assumed in a stylised model. By refusing to deal with it, instead, in the future it will use the downstream monopoly to extract rents from the more efficient upstream rival.³⁷

Comments This discussion shows that it is important to consider the expected evolution of a market when analysing incentives for vertical foreclosure. On the other hand, such an assessment is complicated by the uncertainty that comes with forecasts, and it would be controversial, and arguably inappropriate, if intervention relied on a speculative assessment of possible market developments.

Moreover, the above discussion (and the formal analysis developed in Section 5.3.6) refers to the case in which the rival in the downstream market (in period 1) is a potential entrant. In that case, the incumbent needs to

³⁷ The same logic would apply if there was certain future entry downstream, while only potential entry upstream ahead of that event.

credibly commit to refusal to supply, at least in period 1, so as to exclude the rival. Indeed, if refusal to supply was not credible, the rival would pay the entry cost and would enter the market even if the incumbent engages in it anticipating that, once entry occurred, the incumbent would renege on its refusal and would supply the more efficient independent downstream firm so as to extract (some) of its efficiency gains.

The main economic mechanisms are similar when the downstream market is characterised by network externalities rather than scale economies from the supply-side. In that case, one can show that the incumbent has an incentive to engage in refusal to supply to prevent the rival from selling in period 1. Lack of first-period sales prevents the rival from achieving a critical customer base in the downstream market and to compete effectively vis-à-vis the incumbent. In turn, lack of the rival's success (or expansion) in the downstream market may prevent future entry in the upstream market. However, as discussed in the other chapters of this book, under demand-side scale economies the rival does not need to be a potential entrant. It can be already active in the market, but playing a marginal role. Moreover, under demand-side scale economies, exclusion in the downstream market requires preventing the rival from selling in the first period (and not from paying the entry cost). Hence, credibility of the decision to engage in refusal to supply is not necessary in that case.³⁸

The ownership structure of the entrants is another important feature to take into account to assess the incentives for vertical foreclosure. Indeed, if the entrants are subsidiaries of the same firm (or group), then the entrant vertically integrated firm fully internalises the positive effect that entry in one market exerts on the profitability of the unit active in the vertically related market. This makes the incentive to enter both markets stronger than in the case of stand-alone entrants and makes exclusion less likely.

Finally, one may note the similarity between the mechanism of dynamic foreclosure set out here (and in technical Section 5.3.6 in more detail) and the theory of harm discussed in Chapter 1 (on predation), Chapter 2 (on single-product rebates) and Chapter 3 (on exclusive-dealing contracts) based on scale or scope economies. Put simply, these mechanisms rely on the linkages between time periods, that is, between an incumbent's strategy at present and its effects in future periods. In the case of the models of predation, rebates or exclusive dealing (that is, without necessarily having a vertically integrated incumbent), exclusion may take place by currently

³⁸ This discussion is similar to the one developed in Chapter 4 as part of our review of Carlton and Waldman (2002).

denying rivals the necessary scale so they are not viable in the future. In the case of the model of dynamic vertical foreclosure, competitors or entrants are likewise denied the necessary scale either upstream or downstream at present so that they are not viable in the future. This may make entry at the other level less likely in the future, thereby allowing the incumbent to preserve its monopoly position. Alternatively, it may allow the incumbent to transfer its monopoly power to the other level of the market.³⁹

5.3 Formal Models of Vertical Foreclosure*

In this section we provide some formalisation of the potential foreclosing mechanisms set out at an intuitive level in Section 5.2. In Section 5.3.1 we model the Chicago School argument. In Section 5.3.2 we formalise a model of foreclosure based on the existence of regulation on upstream prices. In Section 5.3.3 we present a formal model of foreclosure based on opportunistic behaviour and the commitment problem. In Section 5.3.4 we discuss models of foreclosure in the presence of uncertainty and risk-averse firms. In Section 5.3.5 we present models of foreclosure in the presence of upstream competition and a raising rival's costs motive by the vertically integrated incumbent. Finally, in Section 5.3.6 we develop a model of vertical foreclosure in a dynamic environment.

As the analysis below will emphasise, in each model some specific assumptions are crucial for the result that vertical foreclosure is a profitable strategy for the incumbent. Our reading of this literature is that vertical foreclosure is possible, but it arises under special circumstances, which need to be identified and assessed with care.

5.3.1 The Chicago School Result*

In this section we set up a baseline model that allows us to show that the vertically integrated incumbent has no incentive to engage in vertical foreclosure when it can extract sufficient rents from the independent downstream firm. Starting from this analysis we will later explore under which conditions imperfect rent extraction leads to vertical foreclosure.

³⁹ In Section 5.6 we review a few cases (for example, *Genzyme*) where some of the industry features set out in those decisions and judgments may be interpreted as being consistent with a theory of harm based on dynamic vertical foreclosure.

5.3.1.1 A simple model with inelastic demand and linear pricing*

Assume an indispensable input is sold by a monopolist seller, U_I , which is the upstream affiliate of the vertically integrated firm I . The final homogeneous product can be sold by I 's downstream affiliate D_I , and by one independent firm, D_R . The latter is assumed to be more efficient: $c_{DR} = 0 < c_{DI} = c$. Production of the input is costless: $c_{UI} = 0$, and there are no fixed costs. Market demand is given by $Q = 1$ if the final price p does not exceed v , with $v > c$. The assumption of rigid demand simplifies the analysis. We will discuss at the end of this section a few additional insights that can be obtained under elastic demand. See Figure 5.1 above for an illustration of the market structure. The game is as follows.

1. The vertically integrated firm decides whether to supply the independent downstream firm or to refuse to supply D_R thereby serving all the market via D_I .
2. If it decides to supply the independent firm, the vertically integrated firm offers a *linear tariff* w to D_R . The independent downstream firm D_R decides whether to accept or reject the offer.
3. D_R and D_I choose the price for the final good.⁴⁰ Consumers place their orders. Then D_R orders the input to satisfy demand, paying accordingly, and transforms one unit of the input into one unit of the final good, which is then provided to consumers.

Lemma 5.1 (*Chicago School result*) *The incumbent finds it optimal to let the independent firm serve the final market and to entirely extract its rents. The incumbent has no incentive to engage in vertical foreclosure.*

Proof. Let us consider the case in which the incumbent decides to supply the independent downstream firm. Let us analyse the last stage of the game in which the independent firm and the affiliate choose the retail prices *given* the wholesale price w offered by the incumbent to D_R . In

⁴⁰ Note that here the downstream affiliate D_I is always active. We shall see in Section 5.3.3 that there are situations in which the vertically integrated firms needs to commit to close down its subsidiary in order to maximise rent extraction. Here such a commitment is not necessary. (Likewise, Section 5.3.1.2 will show that such a commitment is not necessary if one assumes elastic demand and allows for either linear pricing or two part-tariffs.)

taking this decision, the independent firm anticipates that the incumbent will undercut – through the downstream affiliate – any price p such that $p - c > w$, that is, any $p > w + c$. When such condition is satisfied, the margin gained on the final product sold by the downstream affiliate is larger than the margin lost on the input that the incumbent fails to sell to the independent firm. Taking into account that the final price cannot exceed v , for a given w the highest price that the independent firm can set so as to capture the final market is $\hat{p} = \min\{w + c; v\}$. Then, the independent firm will find it profitable to supply the final market as long as $\hat{p} = \min\{w + c; v\} \geq w$. Otherwise, the independent firm will not compete in the downstream market.

In stage 2 the incumbent chooses the optimal wholesale price. It anticipates that if $w \leq v - c$, then the independent firm will supply the final market choosing $p = w + c \leq v$. In this case, the incumbent's profits are $\pi_I = w \leq v - c$ and the profits of the independent firm are $\pi_{D_R} = c$. If $w \in (v - c, v]$, then the independent firm will supply the final market choosing $p = v$ and the incumbent's profits are $\pi_I = w \leq v$. Finally, if $w > v$, the independent firm will not find it profitable to compete in the final market. In this case, the incumbent monopolises the downstream market through its less efficient affiliate and its profits are $\pi_I = v - c$. It is easy to see that the optimal wholesale price is $w^* = v$ which induces the independent firm to serve the final market at the monopoly price $p^* = v$ and allows the incumbent to extract all the rents from the independent firm thereby making the first-best (from the firms' perspective) profits $\pi_I^{novf} = v$. Since production is undertaken by the more efficient downstream firm, such profits are larger than the profits that the incumbent would obtain by engaging in vertical foreclosure and by monopolising the final market through its downstream affiliate, $\pi_I^{vf} = v - c$. ■

5.3.1.2 Discussion*

Here we briefly examine some small variations to the baseline model. First we consider the case where the incumbent's downstream subsidiary is more efficient. Then we discuss the case of elastic demand (and also briefly draw some possible implications for how the incumbent's combination of wholesale and retail prices should be assessed). Finally, we examine the case of two-part pricing.

The efficient firm always produces In the example above, the incumbent's downstream subsidiary is less efficient than the independent downstream firm. Consider instead the case in which it is more efficient than the downstream rival, so that $c_{DR} > c_{DI}$. In the last stage when downstream firms choose retail prices, for given wholesale price w , the incumbent has an incentive to undercut any price p that the downstream rival can profitably offer. Indeed, (weakly) profitable prices for firm D_R are such that $p \geq w + c_{DR}$. Since $c_{DR} > c_{DI}$, one obtains that $p - c_{DI} > p - c_{DR} \geq w$: by slightly undercutting any price that is profitable for D_R the incumbent obtains a margin on the units sold by the downstream affiliate ($p - c_{DI}$) that exceeds the margin lost on the input that the incumbent fails to sell to the downstream rival (w). As a result, the equilibrium price in the final market is $p^* = \min\{w + c_{DR}, v\}$, with the downstream affiliate D_I supplying the whole market, and the vertically integrated incumbent making profits equal to $\pi_I = p^* - c_{DI}$. At stage 2 the incumbent chooses the optimal wholesale price. The own affiliate dominates the final market for any w , but the incumbent maximises profits by choosing $w^* \geq v - c_{DR}$, so that $p^* = v$ and $\pi_I^{novf} = v - c_{DI}$, which are the maximal profits that can be obtained in this market (since the downstream affiliate is more efficient than the independent rival) and correspond to the maximal total welfare (since demand is rigid and final consumers are extracted their entire surplus through the final price v). Note that the independent rival is foreclosed because it is less efficient than the upstream affiliate. Then foreclosure is efficient from the welfare point of view (total welfare would amount to $v - c_{DR} < v - c_{DI}$ if the independent rival produced the final product). One could see the Chicago School result in this perspective: the vertically integrated firm will always have an incentive to make the most efficient downstream firm serve the final market.

Note also that in this situation the vertically integrated incumbent is not sacrificing profits. By considering the wholesale price $w^* = v - c_{DR}$ and the final price $p^* = v$, one obtains $p^* - c_{DI} = v - c_{DI} > w^* = v - c_{DR}$: since the own affiliate is more efficient than the independent rival, the margin that the vertically integrated firm obtains on the units sold downstream exceeds the margin lost on the input sales that it fails to make to the independent rival (see footnote 23).

The case of elastic demand and linear pricing In the inelastic demand example analysed in Section 5.3.1.1, the vertically integrated firm manages to extract all profits from the more efficient downstream rival when it supplies it. Hence, it will have no incentive to engage in vertical foreclosure. But no such incentive would emerge either in a less specific model with elastic demand and linear pricing. By appropriately choosing the wholesale price, the incumbent would obtain profits $\pi_I = \pi^m(c)$, that is, the same profits as in the case in which it monopolises the market through its own downstream affiliate. Rent extraction would be *imperfect* in this case, but the incumbent would *not* have an incentive to exclude the independent downstream firm in this scenario either.⁴¹

To see this, notice that, as shown above, through the choice of the wholesale price w and through the competitive pressure exerted by the own affiliate, the incumbent can indirectly control the choice of the retail price by the independent firm, which will be $p = w + c$.⁴² Hence, the incumbent chooses the optimal wholesale price by solving $\max_w [wD(w + c)]$, which is equivalent to solving $\max_p [(p - c)D(p)]$, where $D(p)$ is the demand function for the final product. The optimal choice $w^* = p^m(c) - c$ induces the downstream firm to choose the retail price $p^* = w^* + c = p^m(c)$: the incumbent makes profits $\pi_I = \pi^m(c)$ – which are the same profits that it would obtain by directly monopolising the final market through the own affiliate – and the more efficient independent firm makes profits $\pi_R = cD(p^m(c))$.

Margin squeeze as exploitative abuse? From this analysis one can also see that, in the case of elastic demand, the optimal pricing policy of the incumbent is such that the difference between the retail price p^* and the incumbent's downstream cost is exactly equal to the wholesale price w^* . In other words, it would pass (by indifference) the test that is typically used to identify margin squeeze. (See footnote 23.) However, it is enough to assume that the independent firm can produce the final product with less than one unit of the input (as opposed to the incumbent's affiliate that uses a one-to-one transformation technology) to show that

⁴¹ Rent extraction is imperfect because the incumbent is constrained to use a linear wholesale price. *A fortiori*, the incumbent would not have an incentive to exclude the more efficient independent firm if it could use two-part tariffs and extract all the rents. See below.

⁴² For the sake of exposition, we are focusing on values of w such that $w + c < p^m(w)$. The result remains valid if one considers also the cases in which such condition is not satisfied.

the optimal incumbent's pricing policy would *strictly* fail such test.⁴³ In this case, however, failure of the test *would not identify exclusion*, since the incumbent finds it optimal to let the independent downstream firm operate and, as shown above, the independent firm makes a positive profit. If a distortion exists, it is because the incumbent chooses a high wholesale price so as to extract rents from the downstream rival. Clearly, a lower wholesale price would support lower retail prices and would improve allocative efficiency. Moreover, it would increase the independent firm's margin, thereby reinforcing incentives to enter the downstream market with new products or more efficient technologies in the long-run. However, it seems to us that sanctioning the incumbent's wholesale price in this case would be similar to intervening through the enforcement of excessive (or unfair) pricing provisions.⁴⁴

The case of elastic demand and two-part pricing In the case of elastic demand and linear pricing, we have seen that foreclosure does not take place even if the incumbent is not able to extract all efficiency rents from the rival. *A fortiori*, the same result holds when the incumbent can rely on more general pricing schemes.

Consider the setting described at the start of Section 5.3.1, but assume that final demand is given by $Q = D(p)$ (satisfying the usual regularity assumptions) and that the upstream firm can use two-part tariffs when trading with the independent downstream firm, namely a price schedule composed by a fixed and a variable component: $t(q_R) = wq_R + T$.

Following the logic of the proof of Lemma 5.1, given the offered tariff $t(q_R)$ the downstream independent firm will set the highest final price that the vertically integrated firm has no incentive to undercut, that is $\hat{p} = \min\{w + c; p^m(w)\}$, where $p^m(w)$ is the price that solves $\max_p[(p - w)D(p)]$. Let us focus on the case in which $w + c \leq p^m(w)$, so that $\hat{p} = w + c$. By setting this price the downstream firm makes profits $\pi_R = (\hat{p} - w)D(\hat{p}) - T = cD(w + c) - T$.

At stage 2 the vertically integrated firm chooses the optimal tariff. Note that not only the choice of the variable component w (combined with the competitive pressure exerted by the own downstream affiliate)

⁴³ See Jullien et al. (2014) for the formal analysis of this case and for an extensive discussion of margin squeeze as an exploitative abuse.

⁴⁴ See Motta and de Streel (2007) for a discussion of the relative merits of such type of intervention.

allows the incumbent to control the retail price set by the independent firm ($\hat{p} = w + c$), but the fixed fee also allows the incumbent to extract the profits made by the independent firm. Indeed, for given w the incumbent will choose $T = cD(w + c)$, so that the independent firm is left with the same payoff as in the case in which it rejects the offer and does not produce. As a result, the incumbent chooses w so as to solve $\max_w[wD(w + c) + T] = \max_w[(w + c)D(w + c)]$. This is equivalent to $\max_p[pD(p)]$: the incumbent's optimal variable component, $w^* = p^m(0) - c$, is chosen in such a way to induce the independent firm to set the retail price $p^* = w^* + c = p^m(0)$ that maximises total industry profits under the use of the independent firm's more efficient technology (recall that $c_{DR} = 0$). The incumbent appropriates such maximal industry profits through the combination of the variable component $w^* = p^m(0) - c$ and of the fixed fee $T = cD(p^m(0))$ and achieves the first-best (from the firms' perspective).⁴⁵ Finally, note that $w^* + c = p^m(0) < p^m(w^*)$, so that the restriction imposed at the beginning of the proof is satisfied. Since the vertically integrated firm obtains the maximal industry profits by trading with the independent firm, it has no incentive to engage in vertical foreclosure.

5.3.2 Regulation of the Wholesale Prices*

A reason why a vertically integrated firm may not be able to extract all rents from the downstream market is that there may be sectoral regulation which prevents the upstream firm from charging the optimal prices to independent downstream firms. In Section 5.3.2.1 we show that this may make it more profitable for the vertically integrated firm to refuse to supply the downstream rivals and to monopolise the downstream market through its less efficient downstream affiliate. We also show that – when it happens – such foreclosure is welfare-detrimental. Of course, a vertically integrated firm that is subject to regulation might be unable to engage in outright refusal to supply and may foreclose the downstream rival in different ways. For instance, it might reduce the quality of the supplied input, it might degrade interconnection or it might claim that it has not enough capacity to serve the rival. Alternatively, it might choose a combination of wholesale and retail prices that make it unprofitable

⁴⁵ An alternative way to achieve the first-best consists of committing not to operate the own affiliate, if possible, combined with a two-part tariff such that $w^* = 0$ and $T = \pi^m(0)$.

for the downstream rivals to operate, that is, it may engage in margin squeeze. We model this possibility in Section 5.3.2.2 and we show that engaging in margin squeeze is more costly than engaging in refusal to supply for the vertically integrated incumbent. Indeed, choosing retail prices in such a way that the more efficient downstream rival is foreclosed entails a sacrifice of profits for the vertically integrated incumbent – similarly to what happens with predatory pricing – and allows it to make profits that are lower than the vertically integrated monopoly profits.

5.3.2.1 Regulation and incentives to engage in refusal to supply*

The model

The model is the same inelastic demand model with linear pricing as in Section 5.3.1, with the independent firm, D_R , being more efficient: $c_{DR} = 0 < c_{DI} = c$. The only difference is that we assume that regulation of the input market imposes a cap \bar{w} on the wholesale price: $w \leq \bar{w} < v$. Figure 5.2 illustrates the market structure. The game is as follows.

1. The vertically integrated firm decides whether to supply or to refuse to supply D_R .
2. If it decides to supply it, the vertically integrated firm offers a *linear tariff* w to D_R , with $w \leq \bar{w} < v$. D_R decides whether to accept or reject the offer.

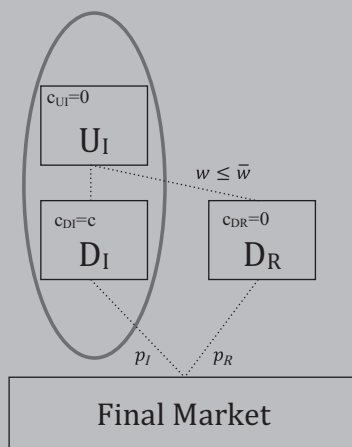


Figure 5.2. Regulated wholesale price.

3. D_R and D_I choose the price for the final good. Consumers place orders. Then D_R orders the input to satisfy demand, paying accordingly, and supplies consumers.

We now show that the vertically integrated firm will find it profitable to foreclose the downstream rival depending on the level of the regulatory cap. Intuitively, the lower the cap, the (weakly) lower the incumbent's profits when supplying the independent rival and the more likely foreclosure prevails. By contrast, the less efficient the downstream affiliate, the lower the profits obtained when it monopolises the downstream market, the less likely foreclosure is. Note that, in this context, foreclosure is always welfare-detrimental when it arises. Qualitatively, results would be the same in a more general model with elastic demand.

Lemma 5.2 (*Refusal to supply by a regulated monopolist*)

- (a) *Refusal to supply is the optimal choice for the vertically integrated firm iff $\bar{w} < v - c$. If instead $\bar{w} \geq v - c$, the independent firm is supplied.*
- (b) *At the equilibrium where refusal to supply occurs, the refusal to supply is welfare-detrimental.*

Proof. (a) If the incumbent decides to supply the independent firm, it is easy to see that it has no incentive to choose a wholesale price strictly below the cap. Indeed, we have seen in the proof of Lemma 5.1 that if $w \leq v - c$, then the independent firm will supply the final market choosing $p = w + c \leq v$, resulting in the incumbent having profits $\pi_I = w \leq v - c$, which are increasing in w . Hence, if $\bar{w} < v - c$, then the incumbent's profits are $\pi_I^{novf} = \bar{w} < v - c$, since the independent firm will supply the final market choosing $p = \bar{w} + c < v$. If $\bar{w} \in [v - c, v)$, then the incumbent's profits are $\pi_I^{novf} = \bar{w} < v$ since the independent firm will supply the final market choosing $p = v$ (recall that D_R will set $\hat{p} = \min\{w + c; v\}$). If the incumbent refuses to supply the independent downstream firm, it will monopolise the final market through the less efficient affiliate, setting the final price $p = v$ and making profits $\pi_I^{vf} = v - c$. Vertical foreclosure turns out to be more profitable than supplying the independent firm if and only if $\bar{w} < v - c$.

(b) Let us focus on the case in which foreclosure arises at equilibrium, that is, in which $\bar{w} < v - c$. When the vertically integrated incumbent refuses to supply, the final price is $p = v$ and total welfare is given by:

$$W^{vf} = CS^{vf} + \pi_I^{vf} + \pi_{DR}^{vf} = 0 + v - c + 0 \quad (5.1)$$

Instead, when the incumbent supplies the independent firm at the wholesale price \bar{w} the final price is $p = \bar{w} + c$ and total welfare is given by:

$$W^{novf} = CS^{novf} + \pi_I^{novf} + \pi_{DR}^{novf} = v - \bar{w} - c + \bar{w} + c = v \quad (5.2)$$

The final price is higher under vertical foreclosure, but in this model with rigid demand this does not entail a deadweight loss. By revealed preferences, when it chooses to engage in vertical foreclosure the incumbent earns higher profits as compared to a case in which it supplies the independent firm. The latter is excluded from the market under vertical foreclosure and suffers a profit loss. Overall, total welfare is lower under vertical foreclosure because production of the final product is undertaken by the less efficient downstream operator. ■

Margin squeeze is not a good substitute for refusal to supply One may wonder whether the incumbent could rely on a margin squeeze rather than on a refusal to supply in the environment analysed in this section.⁴⁶ Indeed, to make the activity of the downstream rival unprofitable, the incumbent should set the highest possible wholesale price, $w = \bar{w}$, and a retail price $p_I = \bar{w}$. The downstream subsidiary would then supply all the market, resulting in a profit $\pi_I^{ms} = \bar{w} - c$, which is strictly lower than the profit $\pi_I^{novf} = \bar{w}$ it would make by supplying the rival.

In an environment with capped access prices, therefore, a margin squeeze does not appear as a good choice for the vertically integrated incumbent, as it would entail a profit sacrifice relative to the case of serving the rival. So if (explicit or implicit) refusal to supply is not

⁴⁶ We acknowledge that it may be rare in practice for there to be a regulated wholesale price, but no duty to deal. On the other hand, even a duty to deal may occasionally be circumvented by a dominant firm. For example, in *Telekomunikacja Polska*, the European Commission found that the eponymous operator (the sole provider of wholesale broadband access and local loop unbundling in Poland at the time) proposed unreasonable conditions in the draft contracts, delayed the negotiations and refused access to (i) its network, (ii) subscriber lines and (iii) reliable and accurate information required by alternative operators. The Commission concluded that, together, these practices prevented alternative operators from competing effectively in the retail market. The General Court fully upheld.

feasible, one can still observe that the vertical integrated incumbent optimally engages in margin squeeze, but the rationale for such a strategy would have to come from a *dynamic* context in which the losses suffered in the early periods (or the profits sacrificed in the early periods) to squeeze the rival's margin can be recouped in later periods. Recoupment will be possible if the vertically integrated incumbent will be able to operate as a monopolist (more generally, will see its market power increase) in later periods. Then, it must be that a squeeze of the rival's margin in early periods leads to exclusion (or marginalisation) of the downstream rival in later periods, for instance due to the existence of scale economies in the downstream market, like in the model developed in Chapter 1 to rationalise predation, or due to the presence of financial frictions, like in the financial theory of predation.

The purpose of the next section is precisely to shed light on the incumbent's incentives to engage in such a predatory-like margin squeeze, by building on the model with scale economies analysed in Chapter 1 with the addition of an upstream market.⁴⁷

5.3.2.2 Regulation and incentives to engage in (exclusionary) margin squeeze*

The model

We assume that there are *two* buyers, B_1 and B_2 , with unit demand and valuation $v = 1$ for a homogeneous product. Two downstream firms compete to supply the buyers: firm D_I and firm D_R . Production of one unit of the final product requires one unit of an input supplied by an upstream monopolist U_I . Production of the input is costless: $c_{UI} = 0$. Firm D_I is vertically integrated with the upstream monopolist and its downstream marginal cost is equal to c . The independent downstream firm D_R incurs a downstream cost f for the production of the first unit of the final product and a downstream cost equal to 0 for the second unit. We assume that:

$$c < f < 2c \quad (5.3)$$

This assumption implies that, given the same input cost, firm D_I is more efficient than the independent rival in the production of a single unit of the final product, whereas the rival is more efficient in the production of two units. Hence, efficiency considerations would require firm D_R to supply the two buyers.

⁴⁷ See also Jullien et al. (2014) for a similar model.

Firms play the following game.

1. First period:

- (a) The upstream monopolist supplies the independent downstream firm D_R at the regulated wholesale price $w_1 \leq \bar{w}$.
- (b) Firms D_I and D_R compete in the downstream market, simultaneously setting prices p_I^1 and p_R^1 to buyer B_1 .
- (c) Buyer B_1 decides from whom to buy. Then D_R , if chosen by the buyer, orders the input to satisfy demand, paying accordingly and transforms one unit of the input into one unit of the final good, which is then provided to B_1 .

2. Second period:

- (a) The upstream monopolist supplies the independent downstream firm D_R at the regulated wholesale price $w_2 \leq \bar{w}$.
- (b) Firms D_I and D_R compete in the downstream market, simultaneously setting prices p_I^2 and p_R^2 to buyer B_2 .
- (c) Buyer B_2 decides from whom to buy. Then D_R , if chosen by the buyer, orders the input to satisfy demand, paying accordingly, and transforms one unit of the input into one unit of the final good, which is then provided to B_2 .

Note that, as discussed in Chapter 1, the simple base model does not apply only to a situation in which the (downstream) incumbent has already sunk a fixed entry cost whereas the independent rival has not, but also to more general situations in which both firms are active in the downstream market but, because of scale economies that the independent rival has not fully exploited yet, the downstream affiliate is more efficient than the rival in the production of a single unit, whereas the opposite is true regarding the production of the two units. In this section we adopt this interpretation that allows us to abstract from the entry decision and to simplify the exposition.

We begin by analysing the case in which the upstream monopolist is free to choose any wholesale price. The analysis is similar to the one developed in Section 5.3.1.1: the upstream monopolist has no incentive to foreclose the activity of the more efficient downstream firm since it can entirely extract the efficiency rents. We then consider the case in which the upstream monopolist is regulated. Since regulation limits rents extraction, the vertically integrated incumbent has an incentive

to exclude the more efficient independent firm. Moreover, if regulation prevents it from refusing to sell the input to the rival, the incumbent has an incentive to engage in margin squeeze in the first period, which entails a downward distortion of the retail price. For this reason, margin squeeze is more costly for the incumbent as compared to refusal to supply.

The unregulated upstream monopolist

Lemma 5.3 *Absent regulation of the wholesale price, the incumbent finds it optimal to let the independent firm serve the final market in both periods and to entirely extract its rents. The incumbent has no incentive to engage in vertical foreclosure.*

Proof. Let us proceed by backward induction and study the decisions taken in the second period. Consider first the case in which the first buyer chose to buy from the downstream affiliate D_I . Let us analyse the last stage of the game in which the independent firm and the affiliate choose the retail prices *given* the wholesale price w_2 offered by the incumbent to D_R . Note that in this case in which the independent firm did not supply B_1 , firm D_I is more efficient than the rival in supplying the second buyer: by assumption (5.3), $c < f + w_2$ for any $w_2 \geq 0$. For this reason the incumbent has the incentive to undercut any price $p \geq f + w_2$ that the independent firm D_R can profitably offer. Indeed, by assumption (5.3), $f > c$. Then, the margin gained by the incumbent on the final product sold by the downstream affiliate is always larger than the margin lost on the input that the incumbent fails to sell to the independent firm: $p - c > p - f \geq w_2$. The equilibrium in the final market will be such that $p^{*2} = \min\{w_2 + f; 1\}$ and the downstream affiliate supplies B_2 . Anticipating this, the incumbent will optimally choose the wholesale price $w_2^* \geq 1 - f$, which allows the own affiliate to secure the second buyer and to set the monopoly price $p^{*2} = 1$. The incumbent earns $\pi_I^2 = 1 - c$ in the second period, while the independent firm D_R makes zero profits.

Consider now the case in which the first buyer chose to buy from the independent firm D_R . Since the independent firm has produced in the first period, in the second period it is more efficient than the incumbent's affiliate in supplying the second buyer: $0 < c$. The analysis developed in Section 5.3.1.1, Lemma 5.1, applies. The optimal wholesale price is $w_2^* = v = 1$ which induces the more efficient independent firm to serve the final market at the monopoly price $p^{*2} = 1$ and allows the incumbent

to extract all the rents from the independent firm, thereby making the first-best profits (from the firms' perspective) $\pi_I^2 = 1$. These profits are larger than the ones that the incumbent would obtain by monopolising the final market through the downstream affiliate, $\pi_I^2 = 1 - c$. Note that the independent firm D_R makes zero profits in the second period also in this case.

Let us move to period 1. Let us study downstream competition, given the wholesale price w_1 offered by the incumbent to D_R . The independent downstream firm anticipates that, irrespective of whether it secures the first buyer, it will make zero profits in the second period. Hence, it is willing to supply the first buyer as long as the price is above the cost to supply the first buyer, that is, $p_R^1 \geq w_1 + f$. Moreover, the independent firm anticipates that the incumbent will undercut – through the downstream affiliate – any price p such that $p > w_1 + 2c$. Indeed, by undercutting the rival's price, the incumbent will obtain the margin $p - c$ on the first buyer, supplied by the own affiliate, and will also supply the second buyer obtaining the margin $1 - c$; instead, if it lets the rival supply the first buyer, the incumbent obtains the wholesale price w_1 in the first period, and the optimal second period wholesale price $w_2^* = v = 1$ in the second period. Undercutting is more profitable iff $p - c + 1 - c > w_1 + 1$, that is, $p > w_1 + 2c$. Taking into account that the final price cannot exceed 1, for a given w_1 the highest price that the independent firm can set so as to capture the first buyer is $\hat{p}^1 = \min\{w_1 + 2c; 1\}$. In sum, the independent firm will find it profitable to supply the first buyer as long as $\hat{p}^1 = \min\{w_1 + 2c; 1\} \geq w_1 + f$. Otherwise, the independent firm will not compete in the downstream market.

Let us analyse the incumbent's choice of the wholesale price w_1 . The incumbent anticipates that if $w_1 \leq 1 - 2c$, then $\hat{p}^1 = w_1 + 2c > w_1 + f$ by assumption (5.3). Hence, the independent firm will supply the final market choosing $p^1 = w_1 + 2c$. In this case the incumbent's total profits are $\pi_I^{1+2} = w_1 + 1 \leq 2(1 - c)$. If $w_1 \in (1 - 2c, 1 - f]$, then the independent firm will supply the final market choosing $p^1 = 1$ and the incumbent's total profits are $\pi_I^{1+2} = w_1 + 1 \leq 2 - f$. Note that, by assumption (5.3), $2 - f > 2(1 - c)$. Finally, if $w_1 > 1 - c$, the independent firm will not find it profitable to compete in the final market. In this case the incumbent monopolises the downstream market through its less efficient affiliate in both periods and its total profits are $\pi_I^{1+2} = 2(1 - c)$. It is easy to see that the optimal wholesale price is $w_1^* = 1 - f$ which induces the

independent firm to serve both buyers at the monopoly price $p^{*1} = 1$ and allows the incumbent to extract all the rents from the independent firm thereby making the first-best (from the firms' perspective) total profits $\pi_I^{1+2, novf} = 2 - f$. Since the independent firm is more efficient than the own affiliate over *total* production, such profits are larger than the profits that the incumbent would obtain by engaging in vertical foreclosure and by monopolising the final market in both periods through the downstream affiliate, $\pi_I^{1+2, vf} = 2(1 - c)$. Note that, by assumption (5.3), $f > c$: the incumbent's downstream affiliate is more efficient than the independent rival in supplying the first buyer. This implies that in the first period the incumbent would earn higher profits by supplying B_1 through the own affiliate rather than by selling the input to firm D_R . However, by letting it supply the first buyer, the incumbent allows the independent firm to reach the efficient scale and to be very efficient in the provision of the product to the second buyer. Such efficiency rents are extracted by the incumbent through the appropriate choice of the wholesale price and, by the assumption that the firm D_R is more efficient than the downstream affiliate over total production, are large enough to dominate the incumbent's profit sacrifice in the first period. ■

The overall result here that foreclosure is feasible for the incumbent, but it would not be profitable, can be interpreted as an extension of the standard Chicago School argument presented in Section 5.3.1.

The regulated upstream monopolist

We now consider the case in which the regulator imposes a cap \bar{w} on the wholesale price set by the incumbent. To ease the exposition, we impose that $\bar{w} \in [0, 1 - f)$. If we did not impose this restriction, we would find that, for \bar{w} large enough, the incumbent would have no incentive to foreclose the activity of the independent downstream firm because, in spite of regulation, the efficiency rents that it can extract from the downstream rival are sufficiently large. This insight already emerged from the analysis of the one-period model of Section 5.3.2.1. Instead, the purpose of this section is to focus on an environment in which regulation, by limiting rent extraction, generates the incentive to exclude the independent downstream rival, and to understand whether in such environment the incumbent has an incentive to engage in margin squeeze.

The analysis will show that margin squeeze involves a profit sacrifice in the short-run: in the early period the vertically integrated incumbent

would find it more profitable to supply the independent rival rather than to engage in margin squeeze. However, the early profit sacrifice can be recouped in later periods when the vertically integrated incumbent can increase prices thanks to the fact that margin squeeze, by denying the rival efficient scale, has made it a poor competitor (or has excluded it from the market). The analysis also shows that margin squeeze is more costly for the incumbent than engaging in refusal to supply. In other words, if despite the regulated environment the incumbent managed not to supply the rival, this would be a more profitable exclusionary strategy than the margin squeeze.

Proposition 5.4 (*Predatory-like margin squeeze in a regulated environment.*)

- (a) *The equilibrium solutions of this two-period game are as follows:*
- (i) *If $f > 3c/2$ then the equilibrium retail prices are $p^{*1} = \bar{w} + f - c$ in the first period and $p^{*2} = \bar{w} + f$ in the second period. The incumbent's downstream affiliate supplies both buyers.*
 - (ii) *If $f \leq 3c/2$ then the equilibrium retail prices are $p^{*1} = \bar{w} + 2c - f$ in the first period and $p^{*2} = \bar{w} + c$ in the second period. The independent downstream firm supplies both buyers.*
- (b) *When the activity of the independent firm is foreclosed (that is, when $f > 3c/2$) the vertically integrated firm sacrifices profits in the first period: the margin it makes on the sales to B_1 falls short of the margin lost on the input that it fails to sell to the independent rival: $p^{*1} - c = \bar{w} + f - 2c < \bar{w}$.*
- (c) *When vertical foreclosure through margin squeeze arises, the incumbent would make larger profits if it could engage in refusal to supply.*
- (d) *Vertical foreclosure is welfare-detrimental when it arises in equilibrium.*

Proof. (a) From the analysis of the unregulated upstream monopoly it is easy to see that the vertically integrated incumbent has no incentive to choose a wholesale price strictly below the cap. Let us proceed by backward induction and study the decisions taken in the second period. Consider first the case in which the first buyer chose to buy from the downstream affiliate D_I . Given the wholesale price \bar{w} that the independent firm must pay to the incumbent, let us analyse the choice

of the retail prices. As already discussed for the unregulated monopolist, in this case in which the independent firm did not supply B_1 , firm D_I is more efficient than the rival in supplying the second buyer: by assumption (5.3), $c < f + \bar{w}$ for any $\bar{w} \geq 0$. For this reason the incumbent has the incentive to undercut any price $p \geq f + \bar{w}$ that the independent firm D_R can profitably offer. The equilibrium in the final market is such that $p^{*2} = \min\{\bar{w} + f; 1\} = \bar{w} + f$ by the assumption that $\bar{w} \in [0, 1 - f)$, and the downstream affiliate supplies B_2 . The incumbent earns $\pi_I^2 = \bar{w} + f - c < 1 - c$ in the second period, while the independent firm D_R makes zero profits.

Let us consider now the case in which the first buyer chose to buy from the independent firm D_R . Since the independent firm has produced in the first period, in the second period it is more efficient than the incumbent's affiliate in supplying the second buyer: $0 < c$. From the analysis developed in Section 5.3.1.1, Lemma 5.1, we know that the independent firm supplies the second buyer setting a price $\hat{p}^2 = \bar{w} + c < 1$ by the assumptions that $\bar{w} \in [0, 1 - f)$ and that $f > c$. The incumbent makes profits $\pi_I^2 = \bar{w}$ and the independent firm makes profits $\pi_{D_R}^2 = c > 0$.

Note an important difference relative to the case of the unregulated monopolist. In that case, when the independent firm supplied the first buyer and, in the second period, turns out to be more efficient than the downstream affiliate, the incumbent manages to extract all the second-period efficiency rents. Hence, the second period profits that the incumbent makes when, in period 1, it was the independent firm that supplied the first buyer are larger than the ones that it makes when it was its own affiliate that supplied the first buyer. Now that regulation limits the rents that the incumbent can extract from the independent firm, the opposite results emerge: the incumbent makes larger second period profits when the own affiliate supplied the first buyer as compared to the case in which the independent firm did: $\bar{w} + f - c > \bar{w}$ by assumption (5.3). This gives the incumbent the incentive to try and 'win' the first buyer. The incumbent will manage to do so when independent firm's efficiency advantage over total production is not too large, as we show next.

In period 1 the independent firm anticipates that it will make profits $\pi_{D_R}^2 = c > 0$ in the second period if it secures the first buyer. Hence, it is willing to supply the first buyer as long as the price is above the cost to supply it less the second period profits, that is, as long as $p_R^1 \geq \bar{w} + f - c$. Moreover, the independent firm anticipates that the incumbent will undercut, through the downstream affiliate, any

price p such that $p \geq \bar{w} + 2c - f$: by undercutting the rival's price the incumbent will obtain the margin $p - c$ on the first buyer and the margin $\bar{w} + f - c$ on the second; by letting the rival supply the first buyer, the incumbent will obtain the wholesale price \bar{w} in both periods. Hence, the independent firm finds it profitable to secure the first buyer as long as $\bar{w} + 2c - f \geq \bar{w} + f - c$, that is, as long as $f \leq 3c/2$. Note that, in this simple model with rigid demand, the lowest retail prices that firm D_R and the incumbent are willing to offer are rescaled by the amount \bar{w} as compared to the lowest price identified in the predation model of Chapter 1. For this reason the threshold level of the fixed cost such that the incumbent wins competition for B_1 and B_2 is exactly the same across the two models. Such an identity has no general validity.

(b) Note that when $f > 3c/2$, then the assumption that $f < 2c$ implies that $p_1^* - c = \bar{w} + f - 2c < \bar{w}$. Then the vertically integrated incumbent is sacrificing profits in the short-run. However, this pricing policy is overall optimal because, by preventing the downstream rival from supplying B_1 , it denies efficient scale to the rival. This makes the rival a poor competitor in the second period, and allows the incumbent to extract large rents from B_2 , thereby recouping the initial profit sacrifice.

(c) When $f > 3c/2$, the incumbent's total profits are lower than the total profits that the incumbent would obtain by engaging in refusal to supply: $\pi_I^{1+2,ms} = 2\bar{w} + 2f - 3c < 2(1 - c) = \pi_I^{1+2,rs}$ by $\bar{w} < 1 - f$. The intuition is the following. We discussed above that the second period profits of the regulated incumbent are larger when the own affiliate serves the first buyer, which gives the incumbent the incentive to try and 'win' the first buyer. If the incumbent could engage in refusal to supply, in order to win the first buyer it would not supply the input to the independent firm and it would then set the monopoly retail price $p^{*1} = 1$. Instead, if it must supply the input at the regulated wholesale price, in order to win the first buyer the incumbent must distort the retail price, which reduces the profits that the incumbent obtains in the first period. The larger the independent firm efficiency advantage (that is, the lower f), the more severe the retail price distortion that the incumbent must introduce, the lower the profits that it obtains by winning the first buyer.

(d) As already shown for the one-period model, the reason vertical foreclosure is welfare-detrimental is that total production over the two periods is undertaken by the less efficient downstream firm. ■

5.3.3 Foreclosure because of Opportunistic Behaviour*

In this section we present a simplified version of Reisinger and Tarantino (2015) which formalises how the very possibility of opportunistic behaviour may lead to (partial) vertical foreclosure. The intuition is that if the vertically integrated firm is unable to commit not to produce via the downstream affiliate, it will be unable to extract the first-best profits from the (more efficient) downstream firm. Therefore, under some circumstances, the incumbent may have an incentive to foreclose, at least partially, the independent downstream firm.

5.3.3.1 The model*

An indispensable input is sold by a monopolist seller, U_I , which is the upstream affiliate of a vertically integrated firm I . There are two firms which can convert the input into the final homogeneous product with a one-to-one relationship: the downstream affiliate D_I of the vertically integrated firm, and one independent firm, D_R .⁴⁸ We assume that the latter is more efficient than the downstream affiliate, having a downstream marginal cost $c_{DR} = 0$, whereas the former has a marginal cost $c_{DI} = c \in (0, 1/3)$. Production of the input is costless: $c_{UI} = 0$. To make it clear that here they do not play any role, let us also assume that there are no fixed costs of production or retail. Finally, assume that market demand is given by $Q = 1 - p$.

Consider the following game.

1. The vertically integrated firm makes a take-it-or-leave-it offer to D_R . It offers a two-part tariff contract $t_R(q_R) = wq_R + T$, where w is the unit wholesale price and T is a fixed fee.
2. The independent downstream firm D_R decides whether to accept or reject the offer. If it rejects the offer, D_R makes zero profits. If it accepts the offer, it commits to honour the accepted tariff.
3. The independent downstream firm D_R (and the downstream affiliate D_I , if active) simultaneously choose the level of output q_i , order the input necessary to produce it and bring production to the market. All payments are made.⁴⁹

⁴⁸ One would obtain similar results also in a setting with multiple independent rivals.

⁴⁹ Downstream firms compete à la Cournot in the final market. In fact, the third stage can be interpreted as the reduced form of a game in which downstream firms order the input first, and then compete in prices under capacity constraints. See Kreps and Scheinkman

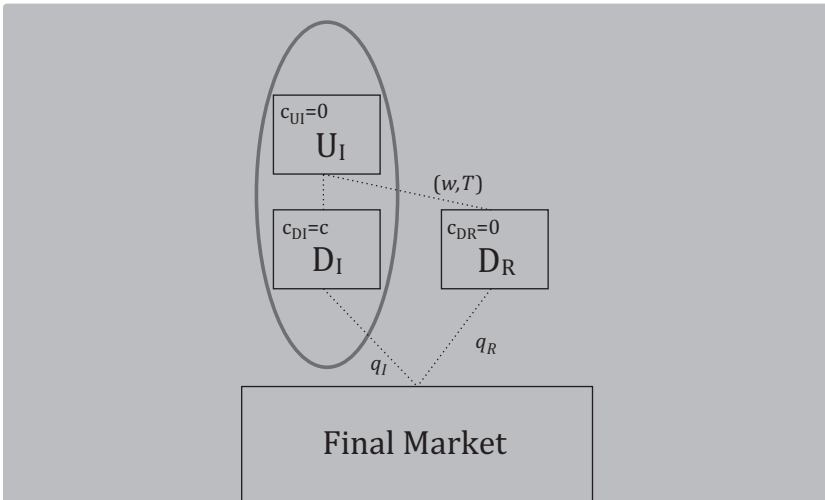


Figure 5.3. The opportunistic behaviour.

The market structure is described in Figure 5.3.

5.3.3.2 Opportunistic behaviour and imperfect rent extraction*

In order to understand the logic of the argument, suppose that the vertically integrated firm could commit not to sell downstream, for instance it could shut down forever D_I , or make a technological commitment for its input to be incompatible with the production process of D_I . In that case, we could reason as if D_I did not exist. It is then easy to see that firm I would be able to earn the first-best profits (from the perspective of firms), that is, the profits it would make if it was vertically integrated with D_R . Firm U_I would offer the intermediate good to D_R at a wholesale unit price equal to its marginal cost, that is, $w = 0$, and a fixed fee $T = 1/4$. Firm D_R would accept (by indifference): at stage 3 it would choose the monopoly output $q_R = 1/2$ and the market clearing price would be $p^m = 1/2$. Then, the independent downstream firm would earn the monopoly profit $1/4$ and would manage to cover the fixed fee $T = 1/4$.⁵⁰

Suppose now that the commitment not to sell via D_I was not possible. Then the outcome identified above cannot be an equilibrium. Firm D_R

(1983). This game refers to a situation in which the downstream firm cannot quickly adjust the purchases of the intermediate input if demand takes unexpected values.

⁵⁰ Note that we obtained the same result when analysing two-part tariffs with Bertrand competition. See Section 5.3.1.

would expect that if it chooses $q_R = 1/2$, then the incumbent would make its downstream affiliate produce $q_I' = \frac{1}{4} - \frac{c}{2} > 0$, where $q_I' = \text{argmax}_{q_I} [(1 - q_I - \frac{1}{2} - c)q_I]$. This choice would allow the incumbent to earn a positive profit in the downstream market which adds to the fixed fee the independent firm has committed to pay. Then, the market clearing price would be $P = 1 - \frac{1}{2} - \frac{1}{4} + \frac{c}{2} = \frac{1}{4} + \frac{c}{2}$, which is lower than the monopoly price $p^m = \frac{1}{2}$ for any $c \in (0, 1/3)$. The independent downstream firm would not earn enough profits to cover the fixed payment $T = 1/4$.

In fact, once it has accepted an offer $t_R(q_R) = wq_R + T$, the independent firm D_R expects that if it chooses the output q_R , then the incumbent would make the downstream affiliate produce the amount $q_I^*(q_R)$, where

$$q_I^*(q_R) = \text{argmax}_{q_I} [(1 - q_R - q_I - c)q_I] = \frac{1 - c - q_R}{2} > 0$$

for any $q_R < 1 - c$. (5.4)

In other words, the independent downstream firm expects that the incumbent will always produce for the final market and will take this into account when deciding on the contract offer. This, in turn, limits the incumbent's ability to extract rents from firm D_R . However, as the following analysis shows, when the independent firm is more efficient than the downstream affiliate, it is not optimal for the incumbent to entirely foreclose its activity. Letting the independent firm produce a positive amount in the final market implies that the good will be produced at lower costs. The incumbent benefits from this more efficient allocation of production since it can extract the profits of the independent firm through the fixed fee.⁵¹ The larger the efficiency gap between the downstream affiliate and the independent firm, the larger the incumbent's benefit from the allocation of some production to the more efficient firm. Hence, the optimal wholesale price decreases with the efficiency gap and the production of the independent firm increases with it. In fact, when the efficiency gap is limited, the incumbent lets the independent firm participate in the final market but supplies it at above its upstream marginal cost, thereby offering less favourable terms than

⁵¹ The benefit of lower production cost dominates the reduction in revenues due to downstream competition.

the ones its downstream affiliate is implicitly supplied at. Instead, when the efficiency gap is large, the incumbent finds it optimal to supply the independent firm at below its upstream cost and offers more favourable terms than the ones implicitly offered to its downstream affiliate.

Proposition 5.5 (Vertical foreclosure in an opportunistic environment.).

(a) In an environment in which there is scope for opportunistic behaviour a vertically integrated incumbent never finds it optimal to engage in **total foreclosure** of an independent downstream firm that is more efficient than the downstream affiliate.

(i) When the efficiency gap is sufficiently small (that is, when $0 < c < 1/5$) it is optimal for the incumbent to engage in partial foreclosure: the independent firm is offered a wholesale price above the incumbent's upstream marginal cost, which is the implicit price at which the downstream affiliate is supplied.

(ii) When the efficiency gap is sufficiently large (that is, when $1/5 \leq c < 1/3$), it is optimal for the incumbent to offer the independent firm a wholesale price below the incumbent's upstream marginal cost.

(b) The incumbent earns lower profits than in the first-best (from the firm's perspective) solution.

(c) Total welfare increases if the vertically integrated incumbent is constrained to decrease the input price charged to the independent downstream firm.

Proof.

a) Given the incumbent's inability to commit not to produce via the downstream affiliate, following an offer $t_R(q_R) = wq_R + T$ the independent firm anticipates that the equilibrium levels of output at stage 3 will be:

$$q_R^*(w, c) = \frac{1 + c - 2w}{3} \quad (5.5)$$

$$q_I^*(c, w) = \frac{1 + w - 2c}{3} \quad (5.6)$$

where w and c indicate the marginal cost of the independent firm and the downstream affiliate, respectively. Hence, given the wholesale price w , firm D_R is willing to accept an offer involving a fixed payment such that:

$$T \leq [P(q_R^* + q_I^*) - w] q_R^* = \frac{(1 + c - 2w)^2}{9} \quad (5.7)$$

The incumbent's optimal choice of the contract to offer to D_R can be written as follows:

$$\begin{aligned} \max_{w,T} [P(q_R^*(w,c) + q_I^*(c,w)) - c] q_I^*(c,w) + wq_R^*(w,c) + T \text{ s.t. } T \\ \leq [P(q_R^* + q_I^*) - w] q_R^* \end{aligned} \quad (5.8)$$

Since the incumbent makes a take-it-or-leave-it offer to firm D_R , it manages to fully extract the profit of the independent downstream firm. Then, the optimisation problem becomes:

$$\max_w [P(q_R^*(w,c) + q_I^*(c,w)) - c] q_I^*(c,w) + P(q_R^* + q_I^*) q_R^*(w,c) \quad (5.9)$$

By substituting $P(Q) = 1 - Q$ and equations identifying the equilibrium levels of output at stage 3, one obtains:

$$\max_w \frac{2(1-c) - w^2 + 5c^2 + w(1-5c)}{9} \quad (5.10)$$

The FOC of the above problem is given by $-2w + 1 - 5c = 0$. Hence the optimal contract entails $w^* = \frac{1-5c}{2}$ and $T^* = 4c^2$. Note that the optimal wholesale price is above the incumbent's upstream marginal cost $c_{UI} = 0$ as long as $c < 1/5$.

b) By offering the optimal contract, the incumbent earns profits $\pi_I^* = \frac{1-2c+5c^2}{4}$. Unless there is no efficiency gap between the independent firm and the downstream affiliate, this profit is larger than the one that the incumbent would obtain by entirely foreclosing the independent firm, but it is smaller than the one that the incumbent would obtain in the first-best solution:

$$\pi_I^{t.f.} = \frac{(1-c)^2}{4} < \pi_I^* = \frac{1-2c+5c^2}{4} < \frac{1}{4} = \pi_I^{f.b.} \text{ for any } c \in \left(0, \frac{1}{3}\right). \quad (5.11)$$

c) Total welfare is given by the sum of the incumbent's profit and consumer surplus. Given w , total welfare W amounts to:

$$\begin{aligned} W &= [P(q_R^*(w,c) + q_I^*(c,w)) - c] q_I^*(c,w) + P(q_R^* + q_I^*) q_R^*(w,c) \\ &\quad + CS(P(q_R^*(w,c) + q_I^*(c,w))); \\ &= \frac{8(1-c) - w^2 + 11c^2 - 2w(1+4c)}{18}. \end{aligned}$$

It is easy to see that total welfare is decreasing in w . ■

5.3.4 Uncertainty and Risk Aversion*

One of the reasons why the upstream monopolist may not be able to extract all the profits from the downstream firms is uncertainty. In what follows, we study a simple model where downstream firms are uncertain about the realisation of costs when they decide on the contractual terms offered by the incumbent and, accordingly, they are uncertain about their profits: the higher their own cost and the lower the rival's, the lower their profits.⁵² We show that if downstream firms are extremely risk-averse, then uncertainty results in the upstream firm leaving large rents to the downstream firms. To save those rents, the upstream firm may prefer engaging in vertical foreclosure and serving the final market through its own (less efficient) downstream affiliate.

This model shows not only that vertical foreclosure may be an optimal response to imperfect rent extraction, but also that vertical foreclosure might be welfare-beneficial. This represents an additional reason to be careful in the analysis of vertical foreclosure.

5.3.4.1 Foreclosure under risk aversion and cost uncertainty*

Similarly to the setting used in the previous sections, assume the upstream affiliate U_I of a vertically integrated firm is a monopolist of the input, and its downstream affiliate D_I produces the output at a constant marginal cost $c < 1/3$.⁵³ We now assume that there are *two* independent downstream firms D_{R1} and D_{R2} , whose marginal cost of production is a random variable: in particular we assume that costs are i.i.d., each firm being a low cost type $c_L = 0$ with probability $r > 0$, and a high cost type $c_H = c$ with probability $(1 - r)$. Note that although stochastic, each independent downstream firm's cost will always be (weakly) lower than D_I 's.⁵⁴

⁵² The model is inspired by Hansen and Motta (2012), which shows that an input manufacturer may want to rely solely on one retailer because when there are several risk-averse retailers competing downstream they will want to be compensated for the risk they incur of facing a more efficient rival. By using one retailer only, instead, this competition risk is avoided. The model presented here extends the intuition to a setting where the upstream monopolist is vertically integrated downstream.

⁵³ The assumption that $c < 1/3$ simplifies the analysis because it ensures that outputs are always positive.

⁵⁴ Since the cost of the downstream affiliate is known, the assumption of two independent downstream firms with private information on their cost is necessary to model the uncertainty created by downstream competition (and subsequent imperfect rent extraction).

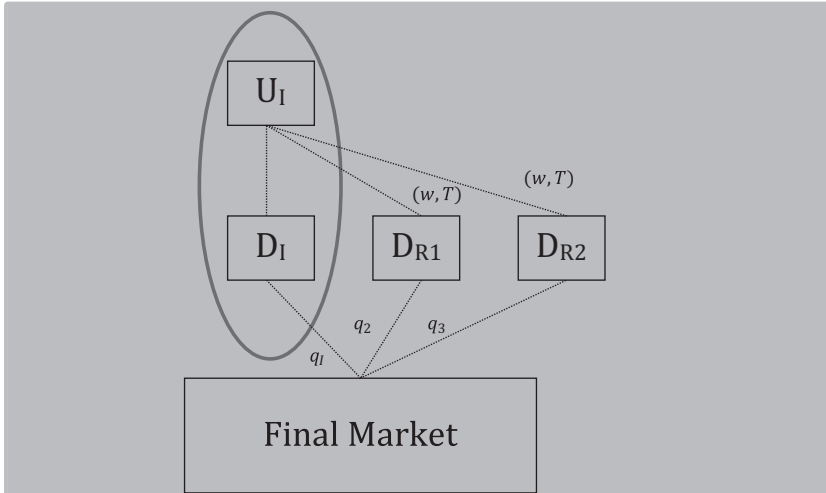


Figure 5.4. Risk aversion.

The upstream firm sets a uniform two-part tariff contract (w, T) , and the two downstream firms have to accept or reject the contract, before the cost realisation. At the moment of choosing their output, though, the cost type has been drawn and it is private information to each firm.⁵⁵ We assume demand to be $p = 1 - Q$ as usual. Also in this case, to focus on a different potential rationale for vertical foreclosure, we assume that the vertically integrated firm can credibly commit not to supply the downstream market through the downstream affiliate D_I . The market structure is illustrated in Figure 5.4.

Crucially for the result of this model, we assume that the vertically integrated firm is risk-neutral but that the independent downstream firms are infinitely risk-averse.

The time of the game is as follows:

1. The vertically integrated firm decides whether to supply the downstream firms or not. If it does, it credibly commits not to sell any output through D_I . Otherwise, it simply refuses to supply the downstream rivals and serves all the market via D_I (vertical foreclosure).
2. If it decides to supply the input, it offers a contract (w, T) .

⁵⁵ Whether the own cost realisation is known at the moment of signing the contract or not does not alter the qualitative results.

3. Independent downstream firms, D_{R1} and D_{R2} , either reject the contract and make zero profits, or accept the contract; in the latter case, they pay T (or credibly commit to do so) and accept to pay w for each unit they buy (but do not order units at this stage).
4. Costs are realised. Each firm only observes its own cost realisation.
5. Each downstream firm chooses its level of output, orders the input, brings production to the market, and all payments are made.⁵⁶

5.3.4.2 A benchmark: risk-neutral downstream firms*

As a benchmark, it may be useful to look first at the case where downstream firms are risk-neutral.

Lemma 5.6 *When downstream firms are risk-neutral, refusal to supply is never more profitable than serving downstream competitors.*

Proof. At the last stage of the game, firm i 's problem ($i = D_{R1}, D_{R2}$) depends on whether it is low cost ($c_L = 0$) or high cost ($c_H = c$):

$$\begin{aligned}\max_{q_{Li}} \pi_{Li} &= \max_{q_{Li}} q_{Li} \{1 - q_{Li} - [rq_{Lj} + (1-r)q_{Hj}] - w\}; \\ \max_{q_{Hi}} \pi_{Hi} &= \max_{q_{Hi}} q_{Hi} \{1 - q_{Hi} - [rq_{Lj} + (1-r)q_{Hj}] - w - c\}.\end{aligned}$$

By taking the FOCs and imposing symmetry ($q_L = q_{Li} = q_{Lj}$ and $q_H = q_{Hi} = q_{Hj}$), we obtain:

$$q_H = \frac{2(1-w) - c(2+r)}{6}; \quad q_L = \frac{2(1-w) + c(1-r)}{6}; \quad (5.12)$$

and the expected profits of a firm of type K at stage 5 are: $\pi_K = (q_K)^2$, with $K = L, H$.

Note, however, that when deciding whether to accept or reject the contract, the downstream firms do not know their type yet; therefore, their expected (gross) profits at stage 3 are:

$$\pi = r\pi_L + (1-r)\pi_H = r \left(\frac{2(1-w) + c(1-r)}{6} \right)^2 + (1-r) \left(\frac{2(1-w) - c(2+r)}{6} \right)^2 \quad (5.13)$$

Therefore, and keeping in mind that here we are considering the benchmark case of risk neutrality, at the first stage of the game the

⁵⁶ The assumption of quantity competition is not inconsequential. When downstream firms compete in prices and products are homogeneous, it turns out that foreclosure does not occur in equilibrium.

vertically integrated firm's problem will be:

$$\max_{w,T} \{2T + 2w(rq_L + (1-r)q_H)\}, \text{ s.t.: } T \leq \pi.$$

By setting optimally $T = \pi$ and solving for the optimal wholesale price, one obtains:

$$w_{rn}^* = \frac{1 - c(1-r)}{4},$$

and the equilibrium variables are:

$$q_H = \frac{1 - c(1+r)}{4}; q_L = \frac{1 + c(1-r)}{4};$$

$$\pi_{I,rn}^{novf} = \frac{(1-c)^2}{4} + \frac{(2-rc)rc}{4}.$$

One may check that in the extreme case in which $r = 1$, that is in which both independent downstream firms are low cost with certainty, $w_{rn} = 1/4$ and $\pi_{I,rn} = 1/4$: by supplying the more efficient independent firms the incumbent extracts the 'first-best' profits, that is, the profits that it would make if it could use the more efficient technology in the downstream market. Of course, supplying the downstream rivals is more profitable than foreclosure, in this case.

In the opposite extreme case in which $r = 0$ and both downstream firms are high cost with certainty, $w_{rn}^* = (1-c)/4$ and $\pi_{I,rn} = (1-c)^2/4$: the incumbent's downstream affiliate and the independent downstream firms are equally efficient; hence, by supplying them, the incumbent extracts the same profits as in the case of vertical foreclosure. In all the intermediate cases in which $r > 0$, supplying the independent firms is more profitable than vertical foreclosure:

$$\pi_I^{vf} = \frac{(1-c)^2}{4} < \pi_{I,rn}^{novf} = \frac{(1-c)^2}{4} + \frac{(2-rc)rc}{4}. \quad (5.14)$$

■

This result can be interpreted as an extension of the standard Chicago School result: it is feasible – but not profitable – for a vertically integrated incumbent to foreclose a more efficient downstream competitor.

5.3.4.3 Infinitely risk-averse downstream firms*

The fact that downstream independent firms are infinitely risk-averse forces the incumbent to leave sufficient rents to them and to distort

the choice of the wholesale price, thereby limiting the profitability of supplying downstream firms. However, the incumbent benefits from supplying downstream rivals that are (weakly) more efficient than its downstream affiliate. The following Lemma shows that, when the probability that the downstream independent firms have a low cost is small, the former effect is strong enough to make foreclosure always optimal. When, instead, the probability that the downstream independent firms are low cost is large enough, then foreclosure is preferred only as long as the incumbent's efficiency gap is not too large.

Lemma 5.7 (a) *When downstream firms are infinitely risk-averse, then refusal to supply is always an optimal choice for the incumbent if the probability that downstream firms have low cost realisations is sufficiently low (that is, if $r \leq \frac{2\sqrt{2}}{3} - \frac{1}{3}$). Otherwise, refusal to supply is an optimal choice if (and only if) $c \leq c^{ina} \equiv \frac{4(1-r)}{5-6r+9r^2}$.*
 (b) *When it arises, refusal to supply is welfare-beneficial.*

Proof. (a) At the last stage of the game, when output is chosen, a firm knows its own cost but not the one of its competitor; since it is infinitely risk-averse, a firm's expected profit corresponds to the worst possible outcome, that is, facing a low cost rival. Accordingly, firm i 's problem ($i = D_{R1}, D_{R2}$) – depending on whether it is low or high cost – will be:

$$\begin{aligned} \max_{q_{Li}} \pi_{Li} &= \max_{q_{Li}} \{q_{Li}(1 - q_{Li} - q_{Lj} - w)\}; \\ \max_{q_{Hi}} \pi_{Hi} &= \max_{q_{Hi}} \{q_{Hi}(1 - q_{Hi} - q_{Lj} - w - c)\}. \end{aligned}$$

By taking the FOCs and imposing symmetry we obtain:

$$\hat{q}_H = \frac{2(1-w) - 3c}{6}; \hat{q}_L = \frac{1-w}{3}$$

with the worst expected profits of a firm of type K at stage 5 being: $\hat{\pi}_H = (\hat{q}_H)^2$.

When deciding whether to accept the contract, the downstream firms do not know their type yet; therefore, and taking once again into account that they are infinitely risk-averse, their expected profits correspond to the worst possible outcome, that is, being high cost and facing a low cost rival:

$$\hat{\pi} = \hat{\pi}_H = \left(\frac{2(1-w) - 3c}{6} \right)^2$$

At the first stage of the game the vertically integrated firm's problem will therefore be:

$$\max_{w,T} \pi_I = 2T + 2w(r\hat{q}_L + (1-r)\hat{q}_H), \text{ s.to: } T \leq \hat{\pi}_H.$$

Optimisation entails setting $T = \hat{\pi}_H$. After replacing and solving the FOCs one obtains:

$$w_{ira} = \frac{2 - 3c(1 - 3r)}{8},$$

and, after replacing:

$$\begin{aligned} \hat{q}_{H,ira} &= \frac{2 - 3c(1 + r)}{8}; \hat{q}_{L,ira} = \frac{2 + c(1 - 3r)}{8}; \\ \pi_{I,ira}^{nf} &= \frac{4 - 4c(3 - r) + c^2(9 - 6r + 9r^2)}{16} \end{aligned}$$

Refusal to supply will therefore be optimal if (and only if):

$$\pi_I^{vf} - \pi_{I,ira}^{novf} = \frac{c[4(1 - r) - c(5 - 6r + 9r^2)]}{16} \geq 0,$$

which holds for:

$$c \leq c^{ira} \equiv \frac{4(1 - r)}{5 - 6r + 9r^2}.$$

Note that the threshold c^{ira} is above $1/3$ when r is sufficiently low, that is, when $r \leq \frac{2\sqrt{2}}{3} - \frac{1}{3}$. In this case, foreclosure is optimal for any feasible value of c .

(b) By revealed preferences, when refusal to supply arises, the incumbent benefits from it. From an *ex ante* perspective, the independent downstream firms are indifferent between foreclosure and no foreclosure: absent foreclosure the profits they expect to make in the final market are extracted through the fixed fee. Consumer surplus under foreclosure is given by $CS^{vf} = (1 - c)^2/8$, whereas expected consumer surplus absent refusal to supply amounts to:

$$\begin{aligned} CS^{novf} &= r^2 \frac{(2\hat{q}_{L,ira})^2}{2} + (1 - r)^2 \frac{(2\hat{q}_{H,ira})^2}{2} + 2r(1 - r) \frac{(\hat{q}_{L,ira} + \hat{q}_{H,ira})^2}{2} \\ &= \frac{4 + c^2(9 + 2r - 7r^2) - 4c(3 - r)}{32} \end{aligned}$$

The difference in consumer surplus between no refusal to supply and refusal to supply amounts to:

$$\begin{aligned} CS^{novf} - CS^{vf} &= \frac{4 + c^2(9 + 2r - 7r^2) - 4c(3 - r)}{32} - \frac{(1 - c)^2}{8} \\ &= \frac{c(1 - r)[c(5 + 7r) - 4]}{32} < 0 \end{aligned}$$

for the feasible value of $c < 1/3$. Since consumers are better off when foreclosure arises, total welfare will be always larger under foreclosure (given we saw the incumbent is better off, by revealed preferences, and the independent downstream firms are indifferent). The intuition is that under strategic substitutes, a risk-averse firm expecting the rival to be low cost, will restrict production, especially when it is high cost. In fact, total production when both firms are high cost is lower than monopoly production under vertical foreclosure. This is the source of consumer loss absent foreclosure. ■

5.3.5 Upstream Competition and the Raising Rivals' Costs Argument*

In this section, we analyse models investigating what may happen when the incumbent faces upstream competition. Under certain circumstances, the incentive to deny the input to independent downstream firms may come from the so-called raising rivals' costs ('RRC') argument as set out in Ordober et al. (1990): the incumbent's withdrawal from the wholesale market increases the price that the downstream rivals have to pay for the input.⁵⁷ In turn, this increases their costs when competing with the incumbent's downstream affiliate, which may be able to increase prices and profits to the detriment of consumers.

In equilibrium, 'foreclosure' does not mean that the downstream competitor is excluded; but this faces endogenously higher costs (higher input prices are an example of 'partial foreclosure'), so that final consumer prices rise and consumer welfare falls, with respect to an outcome absent RRC.

⁵⁷ This requires that there exists some market power upstream. In the absence of upstream market power, the RRC argument is unlikely to apply.

5.3.5.1 The model*

As in the models analysed above, we consider an upstream seller U_I and a downstream seller D_I that belong to the same (vertically integrated) firm I , and an independent rival in the downstream market, firm D_R . Differently from the above analysis, however, we now assume that there is also an independent upstream rival, firm U_R . All firms have zero production costs (except for input prices that downstream firm D_R may have to pay). The upstream input sold by firm U_I and firm U_R is homogeneous. There is a one-to-one input-output transformation technology. Another feature of the model is that, in the downstream market, there is (exogenous) product differentiation.⁵⁸ In particular, firm i 's final demand is:

$$q_i(p_i, p_j) = \frac{1 - \gamma - p_i + \gamma p_j}{1 - \gamma^2}, \quad i, j = I, R; i \neq j, \quad (5.15)$$

with the parameter $0 < \gamma < 1$ measuring the degree of product homogeneity (that is, maximum differentiation arises as γ approaches 0).⁵⁹ Finally, firms compete in prices and are restricted to linear tariffs (prices) both upstream and downstream.

The market structure is described by Figure 5.5.

The timing of the game is as follows:

1. The incumbent chooses (at zero cost) whether to commit *not* to supply any input to D_R . If the incumbent does not commit, it is free to deal with D_R as it wishes. All parties observe (at zero cost) whether the incumbent has committed not to supply D_R .
2. If a commitment has not been made, firm I and firm U_R simultaneously set (linear) wholesale prices w_I and w_R (respectively) to D_R . Otherwise only firm U_R sets price w_R to D_R .
3. Firm D_R chooses the input supplier.
4. Firm D_I and firm D_R set final prices p_I and p_R simultaneously. Consumers place their orders. Then D_R orders the input to satisfy demand, paying accordingly and transforms one unit of the input into one unit of the final product which is provided to consumers.

⁵⁸ Product differentiation is crucial because, as we will see below, it allows both downstream firms to sell a positive amount even in the presence of cost asymmetries.

⁵⁹ It can be shown that these direct demands for the differentiated products can be obtained from the utility function $U(q_I, q_R) = q_I + q_R - (1/2)(q_I^2 + q_R^2 + \gamma q_I q_R)$.

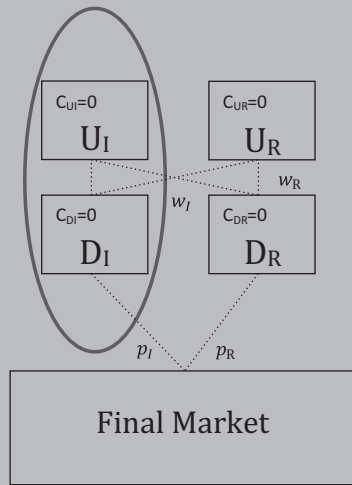


Figure 5.5. Raising rivals' costs.

5.3.5.2 Equilibrium solution*

We solve the game backwards. At stage 4, firm D_I and the downstream rival D_R compete for final consumers. The former has zero cost of production while the latter has a unit cost of w (that is, the input price). The two firms' programmes are $\max_{p_I \geq 0} \{p_I q_I(p_I, p_R)\}$ and $\max_{p_R \geq 0} \{(p_R - w)q_R(p_I, p_R)\}$.

By taking the FOCs and solving one obtains:

$$p_I^* = \frac{2 - \gamma - \gamma^2 + \gamma w}{4 - \gamma^2}; \quad p_R^* = \frac{2 - \gamma - \gamma^2 + 2w}{4 - \gamma^2}; \quad (5.16)$$

$$q_I^* = \frac{2 - \gamma - \gamma^2 + \gamma w}{(4 - \gamma^2)(1 - \gamma^2)}; \quad q_R^* = \frac{2 - \gamma - \gamma^2 - (2 - \gamma^2)w}{(4 - \gamma^2)(1 - \gamma^2)}; \quad (5.17)$$

$$\pi_I^* = \frac{(2 - \gamma - \gamma^2 + \gamma w)^2}{(4 - \gamma^2)^2 (1 - \gamma^2)}; \quad \pi_R^* = \frac{(2 - \gamma - \gamma^2 - (2 - \gamma^2)w)^2}{(4 - \gamma^2)^2 (1 - \gamma^2)} \quad (5.18)$$

While the last stage is common to both cases of foreclosure and competition (by 'competition', we refer to the case in which there has been no commitment by the incumbent not to supply the downstream rival, that is, there is upstream competition), we need to treat the two cases differently when we move backwards.

Competition: the incumbent *has not* committed not to supply the downstream rival

At stage 3, firm D_R will choose to buy from the firm which offers the lower input price. At stage 2, by the standard Bertrand arguments, firms U_I and U_R will set $w_I^c = w_R^c = 0$. Therefore, the price paid by firm D_R will be $w = 0$. Thus, under competition, the equilibrium values will be given by:

$$p_I^c = p_2^c = \frac{1 - \gamma}{2 - \gamma}; \quad q_I^c = q_2^c = \frac{1}{2 + \gamma - \gamma^2}; \quad (5.19)$$

$$\pi_I^c = \pi_{DR}^c = \frac{1 - \gamma}{(2 - \gamma)^2(1 + \gamma)}; \quad (5.20)$$

Foreclosure: the incumbent *has* committed not to supply the downstream rival

At stage 3, firm D_R accepts the wholesale price w_R set by firm U_R provided that this will allow it to produce a positive output: $w_R \leq \frac{2 - \gamma - \gamma^2}{2 - \gamma^2}$.

At stage 2, firm U_R is a monopoly seller to D_R . It therefore solves the following programme:

$$\begin{aligned} \max_{w_R} w_R q_R^*(.) &= \max_{w_R} \left\{ w_R \frac{2 - \gamma - \gamma^2 - (2 - \gamma^2)w_R}{(4 - \gamma^2)(1 - \gamma^2)} \right\} \\ \text{s.t. } w_R &\leq \frac{2 - \gamma - \gamma^2}{2 - \gamma^2}. \end{aligned} \quad (5.21)$$

The constraint is not binding at the optimum and

$$w_R^{vf} = \frac{2 - \gamma - \gamma^2}{2(2 - \gamma^2)} > 0. \quad (5.22)$$

Plugging this equilibrium input price (w_R^{vf}) into the equilibrium values of stage 4 one obtains:

$$p_I^{vf} = \frac{(4 + \gamma - 2\gamma^2)(1 - \gamma)}{2(2 - \gamma)(2 - \gamma^2)}; \quad p_R^{vf} = \frac{(3 - \gamma^2)(1 - \gamma)}{(2 - \gamma)(2 - \gamma^2)}; \quad (5.23)$$

$$\pi_I^{vf} = \frac{(4 + \gamma - 2\gamma^2)^2(1 - \gamma)}{4(2 - \gamma)^2(2 - \gamma^2)^2(1 + \gamma)}; \quad \pi_{DR}^{vf} = \frac{(1 - \gamma)}{4(2 - \gamma)^2(1 + \gamma)};$$

$$\pi_{UR}^{vf} = \frac{2 - \gamma - \gamma^2}{4(2 - \gamma)(2 - \gamma^2)(1 + \gamma)} \quad (5.24)$$

At stage 1, firm I compares the profits that it can obtain when it signs the commitment contract (π_I^{vf}) and to those when it does not (π_I^c). One can show that:

$$\pi_I^{vf} - \pi_I^c = \frac{\gamma(1-\gamma)(8+\gamma-4\gamma^2)}{4(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)} > 0.$$

Therefore, the incumbent prefers to commit not to supply the independent downstream firm. As a result, consumers will be worse off, the downstream firm D_R will obtain lower profits, firm U_R higher profits, and total welfare will decrease:

Proposition 5.8 (Raising rivals' costs.)

- (a) Refusal to supply is always an optimal choice for the incumbent.
- (b) Refusal to supply is welfare-detrimental.

Proof. First, the following inequalities show that prices are higher and consumer surplus lower under foreclosure.

$$p_I^{vf} - p_I^c = \frac{\gamma(1-\gamma)}{2(2-\gamma)(2-\gamma^2)} > 0;$$

$$p_R^{vf} - p_R^c = \frac{1-\gamma}{(2-\gamma)(2-\gamma^2)} > 0;$$

$$CS^{vf} - CS^c = \frac{20+4\gamma-19\gamma^2-\gamma^3+4\gamma^4}{8(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)} - \frac{1}{4-3\gamma+\gamma^3} < 0.$$

Note also that the downstream rival is worse off and the upstream rival is better off under foreclosure:

$$\pi_{DR}^{vf} - \pi_{DR}^c = -\frac{12-9\gamma^2+5\gamma^3}{4(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)} < 0;$$

$$\pi_{UR}^{vf} - \pi_{UR}^c = \frac{2-\gamma-\gamma^2}{4(2-\gamma)(2-\gamma^2)(1+\gamma)} > 0.$$

Finally, one can check that vertical foreclosure results in a welfare loss:

$$W^{vf} - W^c = -\frac{(1-\gamma)(20-8\gamma-19\gamma^2+4\gamma^3+4\gamma^4)}{8(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)} < 0.$$

■

5.3.5.3 Discussion of the assumptions*

This vertical foreclosure result was obtained under three assumptions: (i) downstream competition occurs in prices; (ii) upstream firms are restricted to offer linear tariffs; and (iii) the integrated firm is able to (publicly and irreversibly) commit not to serve the competing downstream firm. Let us relax each assumption in turn.

Strategic substitutability Assumption (i) implies that downstream actions are strategic complements: a more aggressive choice by D_R will be met by a more aggressive response by D_I and *vice versa*. This implies that when U_R increases the input price to D_R , the latter's best reply function will shift upwards, so D_R 's price will increase, which will be met by a higher price by D_I . As a result, competition downstream is relaxed and profits will increase, to the benefit of both upstream firms.

If actions were *strategic substitutes* instead (which would be the case under quantity competition in our simple linear model), a more aggressive choice by one firm would be met by a less aggressive action by the rival, and *vice versa*. Therefore, when – following vertical foreclosure – firm U_R increases the price of the input to D_R , the latter will behave less aggressively and reduce its output, leading to a higher output by downstream affiliate D_I . At the new equilibrium, D_I will have a higher market share and D_R a lower one, determining an increase of the vertically integrated firm's profits. Hence, vertical foreclosure will still be optimal. Therefore, assumption (i) is not crucial for the result.

Two-part tariffs Let us now relax assumption (ii) and consider the possibility for upstream firms to charge two-part tariffs, $T + wq$.

The case in which the integrated firm supplies the downstream rival is identical to the one obtained under linear tariffs. Indeed, when the upstream firms compete for the provision of the input, Bertrand competition will push them to offer $T = w = 0$, so the equilibrium is unchanged.

Let us now consider the case of refusal to supply. The last stage of the game is unchanged. Likewise, the upstream affiliate will continue to provide the input at zero cost to the downstream affiliate.⁶⁰ As for U_R , it

⁶⁰ There is no strategic effect to manipulate internal input prices because internal input prices are non-observable to rivals. Instead, the tariff between the non-integrated firms is observable, and for this reason its choice exerts a strategic effect.

will want to maximise $w_R q_R^*(w_R) + T$, subject to $T \leq \pi_R^*(w_R)$. Hence, its programme will be:

$$\max_{w_R} [w_R q_R^*(w_R) + \pi_R^*(w_R)] = w_R \frac{2-\gamma-\gamma^2-(2-\gamma^2)w_R}{(4-\gamma^2)(1-\gamma^2)} + \frac{(2-\gamma-\gamma^2-(2-\gamma^2)w_R)^2}{(4-\gamma^2)^2(1-\gamma^2)}.$$

By taking the FOCs and solving one obtains:

$$w_R^{vf} = \gamma^2 \frac{(2-\gamma-\gamma^2)}{4(2-\gamma^2)} > 0.$$

By substitution, one obtains:

$$p_I^{vf} = \frac{(4+2\gamma-\gamma^2)(1-\gamma)}{4(2-\gamma^2)}; \quad p_R^{vf} = \frac{(2+\gamma)(1-\gamma)}{2(2-\gamma^2)}; \quad (5.25)$$

$$\pi_I^{vf} = \frac{(4+2\gamma-\gamma^2)^2(1-\gamma)}{16(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)}; \quad \pi_{DR}^{vf} = \frac{(1-\gamma)(2+\gamma)^2}{16(1+\gamma)};$$

$$\pi_{UR}^{vf} = \frac{(1-\gamma)(2+\gamma)^2}{8(2-\gamma)(2-\gamma^2)(1+\gamma)} \quad (5.26)$$

Next, one can check that at stage 1, firm I has higher profits when it commits to foreclose (π_I^{vf}) than when it does not (π_I^c):

$$\pi_I^{vf} - \pi_I^c = \frac{\gamma^3(1-\gamma)(16-8\gamma^2+\gamma^3)}{16(2-\gamma)^2(2-\gamma^2)^2(1+\gamma)} > 0.$$

Therefore, the result that the integrated firm has an incentive to foreclose does not depend on whether contracts are linear or two-part tariffs.⁶¹

Inability to credibly commit not to supply Instead, assumption (iii) is crucial to obtain the results of the model presented in this section. If firm I could revise its decision not to supply firm D_R , it would have an incentive to undercut firm U_R . In other words, absent a credible commitment not to supply D_R , the upstream affiliate of the integrated firm would compete in input supply and the input price paid by firm D_R would not be higher. Nevertheless, there may be situations where it would be realistic for an integrated firm to be able to commit (for

⁶¹ Ordober et al. (1990) also look at possible counterstrategies of the unintegrated rivals facing a vertically integrated firm. The results they obtain in that more general framework hold only under linear prices.

example, technologically) never to supply independent downstream firms.

5.3.6 A Dynamic Model of Vertical Foreclosure to Protect or Transfer Monopoly Power*

In this section, we propose a simple model of vertical foreclosure which is largely inspired by Carlton and Waldman's 'dynamic leverage' model of exclusionary tying due to Fumagalli and Motta (2017b) (we presented it at an intuitive level in Section 5.2.4). However, we shall show that this model goes beyond Carlton and Waldman's motive for foreclosure. Indeed, vertical foreclosure may occur not only to 'protect' a monopolistic position at one level of the vertical chain (whether it is upstream or downstream does not matter), but also to 'transfer' monopoly power from one level to the other. In particular, we show that when entry by an upstream rival cannot be prevented, by eliminating downstream rivals a vertically integrated incumbent may be able to 'create' a downstream monopoly and raise profits.

5.3.6.1 The model*

When the game starts, the vertically integrated incumbent (denoted as firm I) is in the market. Both the upstream and the downstream affiliates of the incumbent have already incurred any fixed entry costs. We consider a game where a downstream rival D_R considers entry immediately, that is, in period 1, while an upstream competitor U_R can enter in a subsequent period, that is, in period 2 (note that the two entrants are not necessarily vertically integrated).⁶²

Upstream firms and (respectively) downstream firms sell perfectly homogeneous inputs and (respectively) outputs. We assume for simplicity that final consumer demand takes a linear form with unit intercept, $Q = 1 - p$. We also assume that potential entrants both upstream and downstream are more efficient than the incumbent: $0 = c_{UR} = c_{DR} = 0$ and $c_{UI} = c_{DI} = c \in \left(\frac{5-2\sqrt{2}}{17}, \frac{1}{4}\right)$.⁶³ However, entrants have to pay a fixed

⁶² One would obtain similar results if immediate entry was feasible in the upstream market, whereas entry was feasible in the downstream market in a later period. See Proposition 5.14.

⁶³ The assumption that the incumbent's affiliates have equal marginal costs simplifies the exposition. The assumption that $c < 1/4$ ensures that $p^m(0) > 2c$ – that is, that the monopoly price having zero marginal costs is higher than the marginal cost of the less

entry cost, respectively, f_U and f_D . Fixed costs need to satisfy the following restrictions:

$$f_U < c(1 - 2c) \quad (5.27)$$

$$c(1 - 2c) - \frac{(1 - c)^2}{8} < f_D < c(1 - 2c) + \frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8} \quad (5.28)$$

Note that we do not impose any lower bound on the fixed costs in the upstream market; indeed, it may be the case that no fixed cost has to be paid to enter upstream. Instead, it is crucial for our mechanism to work that fixed costs in the downstream market are sufficiently large.

Market structure is depicted in Figure 5.6.

The timing of the game is as follows:

1. Period 1, stage 1: The incumbent decides whether to commit to 'refuse to deal' with or to supply the downstream rival.⁶⁴
2. Period 1, stage 2: Firm D_R decides whether to enter (and pay fixed sunk cost f_D) or not;
3. Period 1, stage 3: If D_R is active, with probability 1/2, the incumbent makes an offer to D_R . It offers the contract $t_R(q_R) = wq_R + T$. With probability 1/2, it is D_R that makes an offer to the incumbent. We assume that the incumbent can credibly commit not to operate the downstream unit.⁶⁵
4. Period 1, stage 4: If D_R is active, the contract offer is accepted or rejected. Then active downstream firms choose final prices p_R and p_I , firm D_R orders the input to satisfy demand, paying accordingly, and transforms one unit of the input into one unit of the final product.
5. Period 2, stage 1: Firm U_R decides whether it wants to enter the upstream market (and pay f_U); D_R can still enter if it did not enter in period 1.
6. Period 2, stage 2: With probability 1/2 active upstream firms make offers; with probability 1/2 active downstream firms do.

efficient vertical integrated firm – and that $p_I = p_2 = 2c$ is an equilibrium in the final market when one firm has zero marginal cost and the other marginal cost equal to $2c$. Finally, the assumption $c > \frac{5-2\sqrt{2}}{17}$ ensures that both the new entrants make positive profits at the equilibrium in which entry occurs both upstream and downstream.

⁶⁴ The incumbent's decision to engage in refusal to deal needs to be irreversible, at least in period 1, so as to have a strategic effect.

⁶⁵ This assumption rules out the incumbent's opportunistic behaviour, and together with the absence of regulation, of uncertainty and of upstream competition in the first period,

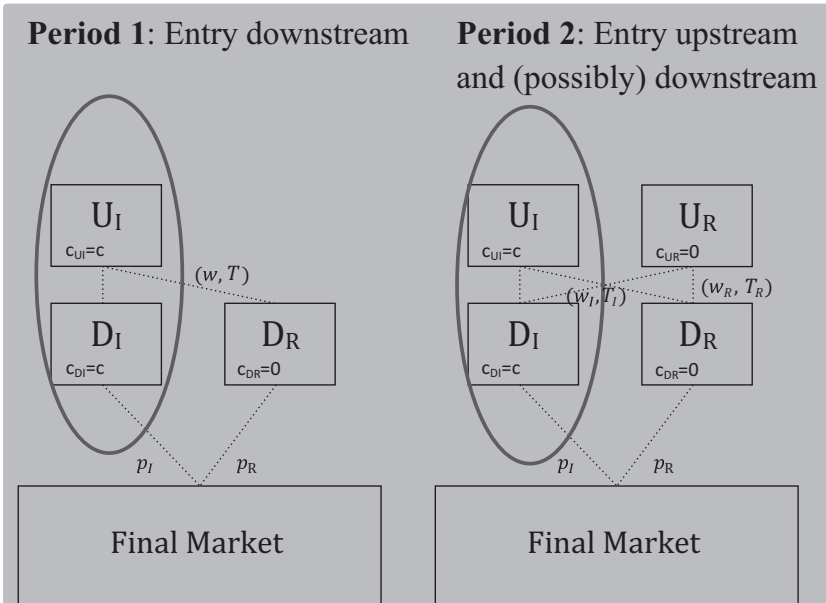


Figure 5.6. Dynamic vertical foreclosure.

7. Period 2, stage 3: Contract offers are accepted or rejected. The active downstream firms set final prices p_I and p_R , orders are made, payments take place and payoffs are realised.

The assumption that upstream firms and downstream firms have equal bargaining power in the negotiation for contract terms simplifies the analysis without loss of generality. For the result it is crucial to exclude only the case in which all the bargaining power is upstream, that is, the case in which the probability that the downstream firm makes the offer is zero. In that case, the downstream rival would obtain zero profits in the first period, if it entered the market. Vertical foreclosure in the first period could not therefore be motivated by the intent to deprive the rival of the profits that are key to cover fixed costs and make entry profitable.

We solve the game by backward induction and we start from period 2 bargaining over input prices, following a decision of the incumbent not to engage in refusal to deal.

allows us to identify a new rationale – other than imperfect rent extraction – for vertical foreclosure.

5.3.6.2 Solution of the model*

No refusal to deal

We now compute the post-entry payoffs in period 2, depending on the configuration of active firms.

(1) Only the independent downstream firm is active

Upstream bargaining power. The incumbent offers firm D_R the contract involving $w = c$ and $T = \frac{(1-c)^2}{4}$ and commits not to sell through its downstream affiliate. Since there is no scope for opportunistic behaviour, firm D_R accepts the contract and the incumbent extracts all the rents from the more efficient downstream competitor.

Downstream bargaining power. Firm D_R offers the incumbent to pay the wholesale price $w = c$ for the input and to pay the fee $T = \frac{(1-2c)^2}{4}$ under the commitment that the incumbent will not sell through its own downstream affiliate. The incumbent accepts the offer. Firm D_R extracts the increase in monopoly profits due to its more efficient production process.

Expected profits of the incumbent and the downstream rival are the following:

$$\pi_I = \frac{1}{2} \left[\frac{(1-c)^2}{4} \right] + \frac{1}{2} \left[\frac{(1-2c)^2}{4} \right]; \quad \pi_{D_R} = \frac{1}{2} \left[\frac{(1-c)^2}{4} - \frac{(1-2c)^2}{4} \right];$$

$$\pi_{U_R} = 0 \quad (5.29)$$

(2) Only the independent upstream firm is active

Upstream bargaining power. Firm U_R offers the incumbent a contract involving $w = 0$ and $T = \frac{(1-c)^2}{4} - \frac{(1-2c)^2}{4}$. The incumbent accepts the offer. Firm U_R extracts the increase in monopoly profits due to the use of its cheaper input.

Downstream bargaining power. The incumbent offers firm U_R to pay the wholesale price $w = 0$ for the input. U_R accepts.

Expected profits of the incumbent and the upstream rival are the following:

$$\pi_I = \frac{1}{2} \left[\frac{(1-2c)^2}{4} \right] + \frac{1}{2} \left[\frac{(1-c)^2}{4} \right]; \quad \pi_{U_R} = \frac{1}{2} \left[\frac{(1-c)^2}{4} - \frac{(1-2c)^2}{4} \right];$$

$$\pi_{D_R} = 0 \quad (5.30)$$

(3) Both upstream and downstream independent firms are active

Upstream bargaining power. The best offer that the incumbent can make to firm D_R involves charging the price $w = c$ under the commitment not to operate the downstream affiliate. This would allow firm D_R to earn $\pi_{D_R} = (1 - c)^2/4$. The upstream rival can match this offer with the contract $w = 0$ and $T = 2c(1 - 2c) - (1 - c)^2/4$ (or a contract involving a slightly lower fee).⁶⁶ Firm D_R accepts the contract offered by the upstream entrant. It will compete in the final market having marginal cost equal to zero vis-à-vis the incumbent's affiliate having marginal cost $2c$.

Downstream bargaining power. The best offer that the incumbent can make to firm U_R is to pay the fee $T = (1 - c)^2/4$ and the price $w = 0$, if it is supplied in exclusivity. Firm D_R can match this offer with the contract $w = 0$ and $T = (1 - c)^2/4$ (or a contract involving a slightly higher fee). Firm U_R accepts the contract offered by the downstream entrant. The latter will compete in the final market having marginal cost equal to zero vis-à-vis the incumbent's affiliate having marginal cost $2c$.

Expected profits of the incumbent and the rivals are the following:

$$\pi_I = 0; \pi_{U_R} = \frac{1}{2} \left[\frac{(1 - c)^2}{4} \right] + \frac{1}{2} \left[2c(1 - 2c) - \frac{(1 - c)^2}{4} \right] = c(1 - 2c) = \pi_{D_R}. \quad (5.31)$$

Note that when they are both active, independent firms share evenly the asymmetric Bertrand duopoly profits generated in the final market when one firm having marginal cost equal to zero competes with a firm having marginal cost $2c$.

(4) No independent firm is active

In this case,

$$\pi_I = \frac{(1 - 2c)^2}{4}; \quad \pi_{U_R} = 0 = \pi_{D_R}. \quad (5.32)$$

The independent firms' payoffs in the different configurations, gross of the entry costs, are summarised in Table 5.1.

Note that each independent firm earns larger profits when the independent firm in the vertically related market is active and competition in that market intensifies. Hence, the fact that an independent firm is

⁶⁶ The assumption $c > \frac{5-2\sqrt{2}}{17}$ ensures that $2c(1 - 2c) > (1 - c)^2/4$.

Table 5.1. Post-entry payoffs under no refusal to deal

D_R, U_R	Active	Not Active
Active	$\Pi_{D_R} = c(1 - 2c)$	$\Pi_{D_R} = \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$
	$\Pi_{U_R} = c(1 - 2c)$	$\Pi_{U_R} = 0$
Not Active	$\Pi_{D_R} = 0$	$\Pi_{D_R} = 0$
	$\Pi_{U_R} = \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$	$\Pi_{U_R} = 0$

active in a market facilitates entry in the vertically related market, as the following Lemma indicates.

Lemma 5.9 Entry decisions in Period 2.

If firm D_R entered in period 1, then firm U_R enters in period 2 for any feasible value of the entry cost f_U .

If firm D_R did not enter in period 1, then the continuation equilibria depend on the level of the entry costs f_D and f_U :

- (i) if $c(1 - 2c) - \frac{(1-c)^2}{8} < f_D \leq c(1 - 2c)$ and $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, then the unique continuation equilibrium is such that both the upstream and the downstream firms enter the market;
- (ii) if $c(1 - 2c) < f_D < c(1 - 2c) + \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$ and $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, then the unique continuation equilibrium is such that firm D_R does not enter the downstream market whereas firm U_R enters the upstream market;
- (iii) if $c(1 - 2c) < f_D < c(1 - 2c) + \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$ and $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} < f_U < c(1 - 2c)$, then the unique continuation equilibrium is that neither firm enters the market.
- (iv) if $c(1 - 2c) - \frac{(1-c)^2}{8} < f_D \leq c(1 - 2c)$ and $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} < f_U < c(1 - 2c)$, then both the upstream and the downstream firms entering the market is a continuation equilibrium. A continuation equilibrium in which neither firm enters the market also exists unless $c \in \left[\frac{5}{17} - \frac{2\sqrt{2}}{17}, \frac{2}{7} - \frac{\sqrt{2}}{14} \right)$ and $c(1 - 2c) - \frac{(1-c)^2}{8} < f_D \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$.

Proof. If firm D_R entered the downstream market in period 1, then firm U_R will earn $\pi_{U_R} = c(1 - 2c) - f_U$ if it enters and $\pi_{U_R} = 0$ otherwise. By assumption (5.27), it decides to enter. Consider now the case in

which firm D_R decided not to enter in period 1. In case (i) entering the upstream market is a dominant strategy for firm U_R . The best reply of firm D_R is to enter the downstream market, since $f_D \leq c(1 - 2c)$. In case (ii) entering the upstream market is a dominant strategy for firm U_R , while not entering the downstream market is a dominant strategy for firm D_R . In case (iii) not entering the downstream market is a dominant strategy for firm D_R . The best reply of firm U_R is not to enter. In case (iv), when $c \in \left[\frac{5}{17} - \frac{2\sqrt{2}}{17}, \frac{2}{7} - \frac{\sqrt{2}}{14} \right)$, then $c(1 - 2c) - \frac{(1-c)^2}{8} < \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$. As a consequence, there exist feasible values of f_D , namely $c(1 - 2c) - \frac{(1-c)^2}{8} < f_D \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, such that entering the downstream market is a dominant strategy for firm D_R . It follows that both firms entering the market is the unique continuation equilibrium. Otherwise, entering the market is the best reply to the entry of the potential entrant in the vertically related market, whereas not entering is the best reply to the decision of the potential entrant in the vertically related market not to enter. Hence, two continuation equilibria may arise: either both firms enter the market or firms fail to coordinate and neither of them enters the market. ■

We can now study the entry decision taken by firm D_R in the first period, as illustrated by the following proposition:

Proposition 5.10 No refusal to deal: *If the incumbent does not engage in refusal to deal, then firm D_R enters downstream in the first period and firm U_R enters upstream in the second period.*

Proof. Firm D_R anticipates that, if it enters in period 1, then firm U_R will enter in period 2 and its total profit is $\pi_{D_R}^{1+2} = \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} + c(1 - 2c) - f_D$. Note that Firm D_R also anticipates that, if it decides not to enter in period 1 then, as established by Lemma 5.9, two possibilities may arise: (i) post-entry profits earned in period 2 alone are insufficient to cover the entry cost; in this case, firm D_R does not enter in period 2 either, and its total profit is zero; (ii) post-entry profits earned in period 2 alone are large enough to cover the entry cost; in this case firm D_R enters in period 2 and its total profit is $\pi_{D_R}^{1+2} = 0 + c(1 - 2c) - f_D$. In both cases, firm D_R finds it more profitable to enter in period 1:

$$\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} + c(1 - 2c) - f_D > \max\{0, c(1 - 2c) - f_D\}. \quad (5.33)$$

In fact, by assumption (5.28), the total post-entry profits earned by firm D_R when it enters in period 1 are large enough to cover the entry cost. Then, period-1 entry is more profitable than no entry in either period (case i). Moreover, when entry in period 2 is profitable (case ii), entry in period 1 is more profitable because it allows firm D_R to earn positive profits for an additional period. ■

Refusal to deal

We consider now the subgame which follows a commitment by firm I not to deal with D_R . This commitment may change the payoffs obtained by the independent firms in the possible configurations of active firms.

(1) Only the independent downstream firm is active

Since D_R cannot trade with the incumbent, its payoff is always zero.

$$\pi_I = \frac{(1-2c)^2}{4}; \pi_{D_R} = 0; \pi_{U_R} = 0 \quad (5.34)$$

(2) Only the independent upstream firm is active

In this configuration, firms' expected payoffs are the same as in the case in which the incumbent does not engage in refusal to deal.

$$\begin{aligned} \pi_I &= \frac{1}{2} \left[\frac{(1-2c)^2}{4} \right] + \frac{1}{2} \left[\frac{(1-c)^2}{4} \right]; \\ \pi_{U_R} &= \frac{1}{2} \left[\frac{(1-c)^2}{4} - \frac{(1-2c)^2}{4} \right]; \pi_{D_R} = 0 \end{aligned} \quad (5.35)$$

(3) Both independent firms are active

Upstream bargaining power. The incumbent cannot trade with D_R . Hence, firm U_R will offer $w = 0$ and $T = 2c(1-2c)$ (or a contract involving a slightly lower fee). Firm D_R accepts. It will compete in the final market having marginal cost equal to zero vis-à-vis the incumbent's affiliate having marginal cost $2c$.

Downstream bargaining power. In this case, contract offers are the same as in the case in which the incumbent does not engage in refusal to deal. The best offer the incumbent can make to firm U_R is to pay the fee $T = (1-c)^2/4$ and the price $w = 0$, if it is supplied with exclusivity. Firm D_R can match this offer with the contract $w = 0$ and $T = (1-c)^2/4$ (or a contract involving a slightly higher fee). Firm U_R accepts the contract offered by the downstream entrant. The latter will compete in the final

Table 5.2. *Firms' payoffs (gross of entry costs)*

D_R, U_R	Active	Not Active
Active	$\Pi_{D_R} = c(1 - 2c) - \frac{(1-c)^2}{8}$	$\Pi_{D_R} = 0$
	$\Pi_{U_R} = c(1 - 2c) + \frac{(1-c)^2}{8}$	$\Pi_{U_R} = 0$
Not Active	$\Pi_{D_R} = 0$	$\Pi_{D_R} = 0$
	$\Pi_{U_R} = \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$	$\Pi_{U_R} = 0$

market having marginal cost equal to zero vis-à-vis the incumbent's affiliate having marginal cost $2c$.

Expected profits of the incumbent and the rivals are as follows:

$$\begin{aligned}\pi_I = 0; \pi_{U_R} &= \frac{1}{2} \left[\frac{(1-c)^2}{4} \right] + \frac{1}{2} [2c(1-2c)]; \\ \pi_{D_R} &= \frac{1}{2} \left[2c(1-2c) - \frac{(1-c)^2}{4} \right]\end{aligned}\quad (5.36)$$

(4) No independent firm is active

In this case,

$$\pi_I = \frac{(1-2c)^2}{4}; \quad \pi_{U_R} = 0 = \pi_{D_R}\quad (5.37)$$

Firms' payoffs in the different configurations, gross of the entry costs, are summarised in Table 5.2.

Note that the decision of the incumbent to engage in refusal to deal limits the profits that firm D_R earns if it enters the market. Trivially, when the independent firm in the upstream market is not active, under refusal to deal firm D_R cannot obtain the input and makes zero profits. However firm D_R 's profits are reduced also when the independent firm in the upstream market is active, because refusal to deal by the incumbent removes competition between the upstream supplier and makes firm D_R more dependent on firm U_R . In fact, in such a case the distribution of the duopoly profits obtained in the final market is more favourable to firm U_R .⁶⁷ Then, the decision to engage in refusal to deal affects the entry decisions, as we show in what follows.

Table 5.2 allows us to analyse the second period entry decision.

⁶⁷ Note that, differently from Ordover et al. (1990), refusal to deal benefits firm U_R but not the incumbent. This is because there is no product differentiation in the final market.

Lemma 5.11 *Entry decisions in period 2.*

If firm D_R entered in period 1, then firm U_R enters in period 2. If firm D_R did not enter in period 1, then it does not enter in period 2 either. Firm U_R enters in period 2 if and only if $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$.

Proof. If firm D_R entered the downstream market in period 1, then firm U_R will earn $\pi_{U_R} = c(1-2c) + \frac{(1-c)^2}{8} - f_U$ if it enters and $\pi_{U_R} = 0$ otherwise. By assumption (5.27), it decides to enter. Consider now the case in which firm D_R decided not to enter in period 1. By assumption (5.28) the post-entry profits collected in period 2 alone are insufficient to cover the entry cost, and it is a dominant strategy for firm D_R not to enter in the second period either. Given that firm D_R does not enter in the second period, firm U_R decides to enter if and only if $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} - f_U \geq 0$. ■

We can now study the entry decision taken by firm D_R in the first period, as illustrated by the following proposition:

Proposition 5.12 *Refusal to deal:* If the incumbent engages in refusal to deal, then firm D_R does not enter the downstream market in either period. Firm U_R enters the upstream market in the second period if and only if $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$.

Proof. Firm D_R anticipates that, if it enters in period 1 then firm U_R will enter in period 2 and its total profit will be $\pi_{D_R}^{1+2} = c(1-2c) - \frac{(1-c)^2}{8} - f_D$. If, instead, it does not enter in period 1, then it will not enter in period 2 either and its total profit will be zero. By assumption (5.28), firm D_R finds it more profitable not to enter. ■

Note that the decision to engage in refusal to supply exerts two effects on firm D_R 's post-entry profits. First, it decreases the second period post-entry profits ensuring that they are insufficient to cover the entry costs. Second, refusal to supply prevents firm D_R from the possibility to earn profits in the first period. As a result, entry in period 1 turns out to be unprofitable.

Vertical foreclosure in equilibrium The above analysis shows that refusal to deal would prevent entry of either the upstream rival, or both the upstream and the downstream rival. However, refusal to deal also sacrifices the incumbent's profits in the first period: the incumbent

would extract some of the entrant's efficiency rents if it dealt with it in the first period. Hence, there is a trade-off for the incumbent: in the short-run, refusing to deal is costly; but in the long-run, by preventing entry by at least one firm, this allows the incumbent not to lose all profits (recall that, when entry occurs both upstream and downstream, the incumbent's profit is zero). In the model presented in this section the incumbent always finds it more profitable to engage in vertical foreclosure in equilibrium.

Proposition 5.13 Profitability of vertical foreclosure and welfare effects

- (a) Under the parameters of the model presented in this section, the incumbent always chooses to engage in vertical foreclosure.
- (i) Vertical foreclosure aims at protecting the incumbent's monopoly position in the upstream market when $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} < f_U < c(1-2c)$.
- (ii) Instead, when $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, vertical foreclosures aims at transferring monopoly power from the upstream to the downstream market.
- (b) Vertical foreclosure is welfare-detrimental.

Proof. a) In the absence of vertical foreclosure, downstream entry occurs in the first period and upstream entry in the second period. The incumbent's total profits are $\pi_I^{1+2,novf} = \frac{(1-c)^2}{8} + \frac{(1-2c)^2}{8} + 0$. Under vertical foreclosure, when $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} < f_U < c(1-2c)$ entry occurs neither in the first period nor in the second period. The incumbent's total profits are $\pi_I^{1+2,vf} = \frac{(1-2c)^2}{2}$. Vertical foreclosure reduces the incumbent's profits in the first period, because by allowing downstream entry the incumbent would partially appropriate the rival's efficiency rents; however, vertical foreclosure, by preventing immediate entry of the downstream rival also prevents future entry of the upstream rival, thereby increasing the incumbent's profits in the second period. Given our feasible values of c , $\pi_I^{1+2,vf} > \pi_I^{1+2,novf}$ (that is, vertical foreclosure is more profitable than dealing with rivals). When $f_U \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, under vertical foreclosure downstream entry does not occur, whereas upstream entry occurs in the second period. The incumbent's total

profits are $\pi_I^{1+2,vf} = \frac{(1-2c)^2}{4} + \frac{(1-c)^2}{8} + \frac{(1-2c)^2}{8}$. Note that second period profits under vertical foreclosure are larger when the upstream rival enters upstream, as compared to the case in which it does not. It is easy to see that $\pi_I^{1+2,vf} > \pi_I^{1+2,novf}$.

b) Total welfare in the absence of vertical foreclosure is given by:

$$\begin{aligned} W^{novf} &= \pi_I^{1+2,novf} + \pi_{DR}^{1+2,novf} + \pi_{UR}^{2,novf} + CS^{1+2,novf} \\ &= \frac{(1-c)^2}{8} + \frac{(1-2c)^2}{8} + 0 + \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} \\ &\quad + c(1-2c) - f_D + c(1-2c) - f_U + \\ &\quad + \frac{(1-c)^2}{8} + \frac{(1-2c)^2}{2}. \end{aligned}$$

In case (i), the decision of the incumbent to engage in vertical foreclosure deters entry both upstream and downstream. Hence, total welfare is given by:

$$\begin{aligned} W^{vf} &= \pi_I^{1+2,vf} + CS^{1+2,vf} \\ &= \frac{(1-2c)^2}{4} + \frac{(1-2c)^2}{4} + \frac{(1-2c)^2}{8} + \frac{(1-2c)^2}{8}. \end{aligned}$$

By point a) the incumbent's profit is larger under vertical foreclosure. Instead, consumers suffer from vertical foreclosure as they pay a higher price in both periods: the incumbent's monopoly price $p^m(2c)$ rather than the monopoly price charged by the more efficient downstream independent firm $p^m(c)$ in period 1; the incumbent's monopoly price $p^m(2c)$ rather than the competitive price $2c$ in period 2. It is easy to show that consumers' loss dominates the incumbent's gain. Moreover, vertical foreclosure reduces the profits of the independent firms whose entry is deterred. It follows that vertical foreclosure is welfare-detrimental.

In case (ii), the decision of the incumbent to engage in vertical foreclosure prevents entry in the downstream market, but not in the upstream market. Total welfare is given by:

$$\begin{aligned} W^{vf} &= \pi_I^{1+2,vf} + \pi_{UR}^{2,vf} + CS^{1+2,vf} \\ &= \frac{(1-2c)^2}{4} + \frac{(1-c)^2}{8} + \frac{(1-2c)^2}{8} + \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} - f_U \\ &\quad + \frac{(1-2c)^2}{8} + \frac{(1-c)^2}{8}. \end{aligned}$$

By point a) the incumbent's profit is larger under vertical foreclosure (and the incumbent's gain from foreclosure is larger than in the case in which foreclosure prevents entry in both markets). Consumers suffer from vertical foreclosure as they pay a higher price in both periods: the incumbent's monopoly price $p^m(2c)$ rather than the monopoly price charged by the more efficient downstream independent firm $p^m(c)$ in period 1; the monopoly price $p^m(c)$ (given that the input is supplied by the more efficient upstream independent firm) rather than the competitive price $2c$ in period 2. (Note that consumers' loss is lower than in the case in which foreclosure prevents entry in both markets.) Also in this case consumers' loss dominates the incumbent's gain. Moreover, vertical foreclosure reduces the profits of the independent firms: entry of the downstream firm is deterred; entry of the upstream firm is less profitable when competition at the other level of the vertical chain is reduced. Then, vertical foreclosure is welfare-detrimental. ■

Vertical foreclosure to 'protect' a monopolistic position The motive for this foreclosure equilibrium is very similar to the one proposed by Carlton and Waldman (2002) in the context of exclusionary tying. In their model, tying is costly in the short-run because the incumbent would benefit from the presence of a more efficient producer of the complementary component. Similarly, in the model presented in this section refusal to supply is costly in the short-run, because it prevents the incumbent from extracting part of the efficiency rents generated by the more efficient downstream firm. Therefore, looking at the first period only, one would conclude that the incumbent has no incentive to engage in vertical foreclosure. However, incentives exist if one looks at the long-run, because refusal to supply affects the future market structure. Refusal to supply deprives the entrant of the key profits that are necessary to make entry in the downstream market viable. Lack of downstream entry, in turn, by limiting competition in the downstream market, reduces the post-entry profits of the upstream independent firm. When entry upstream costs are sufficiently large, refusal to supply also discourages future upstream entry and, by protecting the incumbent's monopoly position in the upstream market, guarantees profits that would be lost otherwise.

Note, however, that differently from Carlton and Waldman (2002), in the model presented in this section the entrants do not need to

be vertically integrated for the mechanism to apply. In their model, when scale economies arise from the supply-side, the entrant and the incumbent compete à la Bertrand in the primary market and are equally efficient. Then, a stand-alone entrant would never enter the primary market, irrespective of whether entry has occurred in the complementary market. The reason why entry in the complementary market creates the incentive to enter the primary one is that a firm that sells the complementary product extracts larger profits from such sales if it also enters the primary market. The reason is that entry in the primary market prevents the incumbent from engaging in price squeeze, that is, in a choice of the price of the complementary product below cost compensated by a higher price for the primary product. Hence, the same firm needs to sell both the primary and the complementary products so as to internalise such an externality. In our context, more simply, downstream entry, by intensifying competition in the downstream market, makes upstream entry more profitable. Independent firms need not to be vertically integrated for this externality to be internalised.

Vertical foreclosure to ‘transfer’ monopoly power The interpretation of this foreclosure equilibrium is different from the previous one. Here, future upstream entry cannot be prevented (indeed, this equilibrium is obtained also for $f_U = 0$), but this is not an obstacle for vertical foreclosure. Rather it is the opposite. Future upstream entry is a problem for firm I only if it follows downstream entry so that the vertically integrated firm does not manage to make any profit in the future. However, by refusing to deal with D_R – while incurring a cost in the short-run – the vertically integrated firm creates a safe monopoly downstream. In the case analysed before, lack of downstream entry also discourages future upstream entry and the vertically integrated firm will be forced to purchase the input from its own inefficient affiliate U_I also in the future. In this case, instead, upstream entry occurs even in the absence of downstream entry and the vertically integrated firm will extract the efficiency rents that the independent firm U_R will generate upstream in the future. Then, when upstream entry occurs for sure, refusal to deal with D_R allows the vertically integrated firm to make larger future profits as compared to the case in which future upstream entry is conditional on current downstream entry, and vertical foreclosure is more – not less – likely as compared to that case. In other words, when upstream entry occurs for sure, the incumbent does not engage in

refusal to supply to protect its upstream monopoly power in the future; it does so to transfer its monopoly position from the upstream to the downstream market. For the reasons discussed above, this new rationale for vertical foreclosure cannot arise in a model à la Carlton and Waldman (2002).

5.3.6.3 Extensions*

Entry is possible upstream initially and downstream in the future So far we have assumed that entry in period 1 can occur only in the downstream market, and that upstream entry can occur only in the future. The same mechanism applies if one assumes the symmetric situation in which entry is currently feasible in the upstream market and will be possible in the downstream market in the future. One needs to impose the following restrictions on entry costs:

$$f_D < c(1 - 2c) \quad (5.38)$$

$$c(1 - 2c) - \frac{(1 - c)^2}{8} < f_U < c(1 - 2c) + \frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8}. \quad (5.39)$$

Proposition 5.14 illustrates the result in this context.

Proposition 5.14 Profitability of vertical foreclosure when current entry occurs upstream

The incumbent always chooses to engage in vertical foreclosure.

- (i) *Vertical foreclosure aims at protecting the incumbent's monopoly position in the downstream market when $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} < f_D < c(1 - 2c)$.*
- (ii) *Instead, when $f_D \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$, vertical foreclosures aims at transferring monopoly power from the downstream to the upstream market.*

This Proposition implies that the incentive to engage in vertical foreclosure arises *irrespective* of whether the initial incumbent's monopoly positions exists in the upstream or in the downstream market.⁶⁸ This is not the case for vertical foreclosure caused by inability to commit (either to specific contractual terms or not to operate the own affiliate),

⁶⁸ The proof of this Proposition follows very closely that of the scenario where entry can occur first downstream, then upstream, and is therefore omitted.

as imperfect rent extraction can only occur when the incumbent's monopoly position lies in the upstream market, and competitive externalities arising in the downstream market are not fully internalised.

Short-term commitment to refusal to supply In the above analysis the decision to engage in refusal to supply discourages downstream entry for two reasons: first it prevents the downstream rival from earning profits in period 1; second, by making the downstream rival more dependent on the upstream firm, it reduces its second-period profits.

If the commitment to refusal to supply does not last forever, but is reversible after one period, then the latter effect is muted and vertical foreclosure is less likely to occur at the equilibrium: the interval of downstream entry costs that sustain refusal to supply would reduce to $c(1 - 2c) < f_D < c(1 - 2c) + \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$. For lower values of the entry cost, downstream (and upstream) entry would occur anyway in the second period and the incumbent would not profit from refusal to supply.

Vertically integrated entrants The above analysis considers the case in which the entrants are independent firms. In this section, instead, we assume that the entrants are subsidiaries of the same firm or group. When this is the case, the vertically integrated firm fully internalises the positive effect that entry in one market exerts on the profitability of the unit active in the vertically related market. This makes the incentive to enter both markets stronger than in the case of independent entrants, and exclusion less likely.

In fact, as Proposition 5.19 will show, the incumbent engages in refusal to supply if (and only if) the total entry costs are sufficiently large, within the feasible intervals defined by assumptions (5.27) and (5.28). Moreover, when the incumbent engages in refusal to supply, then entry in neither vertically related market follows. Differently stated, the incumbent never finds it profitable to engage in refusal to supply so as to shift its dominant position from the upstream to the downstream market. The intuition is that if the costs to enter upstream are very low, so that entry in that market would be profitable under a stand-alone basis, then the vertically integrated firm finds it profitable to enter also downstream. This is because the entry decision is driven by the increase

in *joint* profits produced by downstream entry, not only by the profits generated in the downstream market.

Solution of the model

The incumbent did not commit to refusal to supply

When the entrants are vertically integrated, the entry decisions in period 2 are as follows:

Lemma 5.15 *Entry decision in period 2*

When the incumbent did not commit to refusal to supply, then:

- (i) If affiliate D_R entered in period 1, then affiliate U_R enters in period 2 for any feasible value of the upstream entry cost f_U .
- (ii) If affiliate D_R did not enter in period 1, then in period 2 entry occurs in neither market if (and only if) $f_D + f_U > 2c(1 - 2c)$. Otherwise, entry occurs in both markets.

Proof. If affiliate D_R did enter the downstream market in period 1, then the decision to enter upstream in the following period is driven by the comparison between the upstream entry cost and the increase in *total* profits caused by upstream entry:

$$2c(1 - 2c) - \left[\frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8} \right] \geq f_U. \quad (5.40)$$

Since $2c(1 - 2c) - \left[\frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8} \right] > c(1 - 2c)$ when $c < 1/4$, then by assumption (5.27) the above condition is always satisfied for the feasible values of the upstream entry cost.

If affiliate D_R did not enter the downstream market in period 1, then four possible outcomes may arise in period 2: (i) no entry takes place; (ii) entry takes place only downstream; (iii) entry takes place only upstream; (iv) entry takes place in both markets. However, one can show that outcomes (ii) and (iii) cannot arise. To see this, imagine that upstream entry costs are sufficiently low to make upstream entry alone profitable, that is, $f_U \leq \frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8}$. For the vertically integrated firm it is profitable to enter also downstream if and only if the increase in total profits produced by downstream entry dominates the downstream entry cost:

$$2c(1 - 2c) - \left[\frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8} \right] \geq f_D. \quad (5.41)$$

Since $2c(1 - 2c) - \left[\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} \right] > c(1 - 2c) + \left[\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} \right]$ when $c < 1/4$, then by assumption (5.28) the above condition is always satisfied for the feasible values of the downstream entry cost. This implies that it cannot be that entry occurs only in the upstream market.

Similarly, imagine that downstream entry costs are sufficiently low to make downstream entry alone profitable, that is, $f_D \leq \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$. From the proof of Lemma 5.9 we know that this case may arise when $c \in \left[\frac{5}{17} - \frac{2\sqrt{2}}{17}, \frac{2}{7} - \frac{\sqrt{2}}{14} \right]$, because the lower bound on downstream entry costs is below $\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8}$. Then, for the vertically integrated firm it is profitable to enter also upstream because the increase in total profits produced by upstream entry dominates the upstream entry cost:

$$2c(1 - 2c) - \left[\frac{(1 - c)^2}{8} - \frac{(1 - 2c)^2}{8} \right] \geq f_U. \quad (5.42)$$

The above condition is always satisfied for the feasible values of the upstream entry cost because $2c(1 - 2c) - \left[\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} \right] > c(1 - 2c)$ when $c < 1/4$.

Then, either the vertically integrated firm enters both markets – thereby earning total profits equal to $2c(1 - 2c) - f_U - f_D$ – or it enters neither market, thereby making 0 profits. Entry in neither market is the optimal choice if and only if:

$$2c(1 - 2c) < f_D + f_U. \quad (5.43)$$

Note that, given the upper bounds to the feasible values of f_U and f_D established by assumptions (5.27) and (5.28), for condition (5.43) to be satisfied, it is necessary that $f_D > c(1 - 2c)$ and that $f_U > c(1 - 2c) - \left[\frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} \right]$. ■

We can now move to the entry decision in period 1.

Lemma 5.16 *Entry decision in period 1*

When the incumbent did not commit to refusal to supply, then downstream entry always occurs in period 1, followed by upstream entry in period 2.

Proof. If downstream entry occurs in period 1, total profits of the vertically integrated entrants amount to:

$$\pi_{DR+UR} = \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} + 2c(1-2c) - f_U - f_D > \max\{0, 2c(1-2c) - f_D - f_U\} \quad (5.44)$$

By assumptions (5.27) and (5.28), $f_U + f_D < \frac{(1-c)^2}{8} - \frac{(1-2c)^2}{8} + 2c(1-2c)$. Hence, entering downstream in period 1 and upstream in period 2 is more profitable than not entering at all; it is also more profitable than entering in both markets in period 2, because period-1 entry allows to earn profits for one more period. ■

The incumbent committed to refusal to supply

When the entrants are vertically integrated, the entry decisions in period 2 are as follows:

Lemma 5.17 *Entry decision in period 2*

When the incumbent did commit to refusal to supply, then:

- (i) If affiliate D_R entered in period 1, then affiliate U_R enters in period 2 for any feasible value of the upstream entry cost f_U .
- (ii) If affiliate D_R did not enter in period 1, then in period 2 either entry occurs in both markets (iff $f_D + f_U \leq 2c(1-2c)$) or no entry takes place.

Proof. If affiliate D_R did enter the downstream market in period 1, then entering the upstream market in the following period is the only way for the vertically integrated entrant to make positive profits. Upstream entry occurs if (and only if):

$$2c(1-2c) \geq f_U \quad (5.45)$$

The above condition is always satisfied by assumption (5.27).

If affiliate D_R did not enter the downstream market in period 1, then the argument follows the line of the proof of Lemma 5.15 with minor differences. ■

We can now move to the entry decision in period 1.

Lemma 5.18 *Entry decision in period 1*

When the incumbent did commit to refusal to supply, then entry occurs in neither market if (and only if) $f_U + f_D > 2c(1-2c)$.

Proof. If downstream entry occurs in period 1, affiliate D_R makes zero profits in that period. Entry will occur in the upstream market in period

2 and the vertically integrated entrant will earn total profits $\pi_{2+E}^{1+2} = 2c(1 - 2c) - f_D - f_U$. Hence, if $f_D + f_U \leq 2c(1 - 2c)$, entry will occur in both markets. It is irrelevant whether entry occurs downstream in period 1 followed by upstream entry in period 2, or whether it occurs in both markets in period 2 (by assumption the discount factor equals 1). If $f_U + f_D > 2c(1 - 2c)$, then entry occurs in neither market. ■

Decision to engage in refusal to supply

Proposition 5.19 *Vertically Integrated Entrants*

- a) *When the entrants are vertically integrated, the incumbent chooses to engage in refusal to supply if (and only if) $f_D + f_U > 2c(1 - 2c)$.*
- b) *When refusal to supply occurs, it discourages entry in both markets and it is welfare-detrimental.*

Proof. If the incumbent does not engage in refusal to supply, downstream entry will occur in period 1 and upstream entry in period 2. Total profits in this case are:

$$\pi_I^{1+2, novf} = \frac{(1 - c)^2}{8} + \frac{(1 - 2c)^2}{8} + 0. \quad (5.46)$$

If the incumbent does engage in refusal to supply, and entry in both markets is discouraged (that is, if $f_U + f_D > 2c(1 - 2c)$), then the incumbent sacrifices profits in period 1 but it increases profits in period 2. Given our feasible values of c , the latter effect dominates and vertical foreclosure is profitable:

$$\pi_I^{1+2, vf} = \frac{(1 - 2c)^2}{4} + \frac{(1 - 2c)^2}{4} > \pi_I^{1+2, novf}. \quad (5.47)$$

If, instead, entry occurs in both markets even though the incumbent engaged in refusal to supply, then vertical foreclosure cannot be profitable: the incumbent would sacrifice profits in period 1 and would still make zero profits in period 2.

Note that the case in which entry occurs only upstream as a consequence of the incumbent's decision to engage in refusal to supply does not arise when the entrants are vertically integrated.

The welfare analysis is the same as the one developed in the proof of Proposition 5.13 and leads to the conclusion that vertical foreclosure is welfare-detrimental. ■

5.4 From Theory to Practice

The issue of antitrust intervention in the area of vertical foreclosure (for example, refusal to supply or to license, margin squeeze) is a thorny one, perhaps even more so than in the other areas of potentially exclusionary behaviour reviewed in previous chapters. Part of the difficulty also stems from the possible tension between competition law enforcement and the need for the protection of intellectual property rights (which we also discuss in Section 5.5, with reference to key US and EU case-law).

For this reason, we believe that it is appropriate to start a discussion of the policy implications in the area of vertical foreclosure with the reminder that it is important for antitrust intervention not to dampen the incentives to invest and to innovate (Section 5.4.1). We shall then draw some policy implications from the theories discussed in the earlier sections (Section 5.4.2), and make some summarising remarks (Section 5.4.3).

5.4.1 The Need to Preserve the Incentives to Invest and to Innovate

At the core of this chapter are practices whereby a dominant firm decides not to supply a downstream rival, or to make it pay a very high price for access to its input. As stated at the outset of this chapter, we believe that companies should be free to deal with the parties they wish, and equally free to set the contractual terms in the relationships with them. Only in exceptional circumstances should antitrust intervention be warranted. This is also the position taken by the European Commission in its Guidance Paper on its enforcement priorities in the application of the abuse of dominance provisions:⁶⁹

When setting its enforcement priorities, the [European] Commission starts from the position that, generally speaking, any undertaking, whether dominant or not, should have the right to choose its trading partners and to dispose freely of its property. The Commission therefore considers that intervention on competition law grounds requires careful consideration where the application of Article [102] would lead to the imposition of an obligation to supply on the dominant undertaking. The existence of such an obligation – even for a fair remuneration – may undermine undertakings' incentives to invest and innovate and, thereby, possibly harm consumers. The knowledge that they may have a duty to supply against their will may lead dominant undertakings or undertakings who anticipate that they may become dominant – not to invest, or to invest less, in the activity in question. Also, competitors may be tempted to free ride on investments made

⁶⁹ European Commission (2009).

by the dominant undertaking instead of investing themselves. Neither of these consequences would, in the long run, be in the interest of consumers. [...] (Para. 75)

The European Commission therefore recognises the need to preserve the incentives to invest and/or to innovate.

And so does the US Supreme Court, which – in *Trinko* – held:

Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities. (Part III of the Supreme Court's Opinion)

Unlike perhaps other forms of unilateral conduct, some of the efficiencies that are relevant to the practices reviewed in this chapter flow not so much from vertical foreclosure *per se*.⁷⁰ Rather, they may be interpreted as the need to avoid certain *dynamic inefficiencies*. Such inefficiencies would likely be brought about by certain types of intervention aimed at remedying potential foreclosure.

A prime example of this is the dynamic effect that an access remedy may have on the incentives of firms to innovate or invest in the market under scrutiny. From an economic perspective, a government would be foolish to drastically raise taxes specifically on companies that have been successful thanks to an innovation, risky investments or business acumen. Likewise, competition authorities should think carefully before forcing innovators or firms that have heavily invested in an infrastructure or a product to share the benefits from such facilities or innovations with competitors. Otherwise the incentives to innovate or invest in the first place may significantly diminish, with a potential loss in consumer welfare. This is likely to be true not only in the market under scrutiny: by enforcing a duty to deal, a competition authority may remove the incentive to invest or innovate more generally across other markets. This may occur if investors ended up shying away from high-risk, high-reward ventures in the expectation that – conditional on those investments being successful – their competitive advantages or financial rewards gained as a result of the investment or innovation would be curtailed by antitrust intervention.

In this light, a cautious approach to the assessment of allegations of a refusal to deal (or 'constructive' refusal to deal – that is, a price of the input so high that the downstream rival would not purchase it) will be in

⁷⁰ We review some potential objective justifications in Section 5.4.2.2.

the consumers' interest in the long-run. This is likely to favour dynamic competition and as such is likely to protect the incentives to invest, for example, into new products, services, technologies and infrastructure.

The guiding principle that firms should be free to receive proper remuneration from their investments is consistent with the role that, exceptionally, antitrust intervention may have where the dominant firm's essential input is not due mainly or exclusively to the firm's investments, but to other factors.⁷¹ For instance, suppose that the company that owns some upstream network infrastructure is the successor of a formerly State-owned monopolist. Intervention to facilitate network access by downstream rivals may be more reasonable than in a scenario where the vertically integrated incumbent invested in infrastructure from scratch. Similarly, there may be more scope for intervention in instances where a monopolistic position has been maintained also thanks to anti-competitive practices; where a monopolistic position has arisen without being due to substantial investments;⁷² or where the monopolistic position has existed for such a long time that it would be difficult to argue that the company still needs to obtain proper remuneration for its investments. We are not suggesting that in each of these circumstances a competition authority should always grant entrants or downstream rivals access to an incumbent's network or asset, but simply that such considerations should play a role in the assessment of vertical foreclosure cases, so as to balance the need to promote investments and innovation with a possible desire to foster competition also in the short-run.

5.4.2 Policy Implications

In this section, we attempt to draw some policy implications from the economic models (and theories of harm) presented in the earlier sections, in particular by discussing factors that are likely to be important in the assessment of alleged vertical foreclosure cases, be they refusal to supply

⁷¹ Motta and de Stree (2006) similarly argue that excessive pricing actions under antitrust rules should be rare, due to a number of reasons, including because they reduce the possibility for firms to reap the benefits of their investments and innovations. However, they also identify some *necessary* conditions for (exceptional) intervention. One of these requires that the dominant firm's monopolistic or quasi-monopolistic position is due to past exclusive or special rights or past anti-competitive practices.

⁷² See for instance *IMS Health*, briefly described in Section 5.5.2, where the main input at issue – the existence of a 'brick structure' that had become the standard of data organisation in the pharmaceutical sector in Germany – may be seen to have been due more to network effects than to the company's investments.

or margin squeeze cases. We shall also review some possible legitimate justifications for resorting to vertical foreclosure.

5.4.2.1 Some key factors for the assessment of alleged vertical foreclosure cases

Indispensability of the input A first important observation is that, because of the need to preserve incentives to innovate, a competition authority should be very careful before intervening in cases of vertical foreclosure. As a consequence, one needs something more than mere *upstream* dominance to intervene, and in this light it may make sense to require that an input is *essential* as a necessary condition to intervene under antitrust provisions. This could mean that alternative inputs do not exist or are clearly inferior and that the input cannot be easily duplicated; in other words, there must be extremely high and non-transitory barriers to entry in the upstream market.^{73,74} As we discuss in Section 5.5.2, this principle is broadly valid in the EU for refusal to deal cases: a competition authority must show that the input is indispensable as one of the prerequisites for a finding of abuse of dominance through refusal to deal. However, after the *TeliaSonera* judgment, this is not the case under EU law in the context of margin squeeze, where an abuse may be found even in the absence of input indispensability. We note that this is somewhat contradictory (see the analysis of Section 5.3.2.1): a conduct which all else equal raises more competition concerns (refusal to supply) requires under the current case-law a higher hurdle to find an infringement (input indispensability) than a conduct which is less likely to be anti-competitive (margin squeeze). If anything, other things equal, competition authorities should find it more difficult to intervene in margin squeeze cases.

As we argued in Section 5.4.1, one should also take into account that there are situations in which the need to protect investment incentives is arguably less compelling, for example because the control of the essential input is not due, or it is due to a limited extent, to the dominant firm's

⁷³ We presume that when access to an input is subject to regulation it is because that input is essential. Hence, the requisite of indispensability is very likely to be satisfied if access to the upstream input is regulated.

⁷⁴ The raising rivals' costs models reviewed in Section 5.3.5 show that a vertically integrated firm may engage in refusal to supply with the aim of weakening downstream competition and increasing retail prices rather than of excluding downstream rivals. This theory needs neither indispensability nor upstream dominance to be applied. Still, specific conditions are necessary for it to be valid, such as the ability of the vertically integrated firm to credibly commit not to supply independent rivals.

investment. In such cases, the need to preserve the incentives to invest may be less of a priority in the context of antitrust enforcement.⁷⁵

A robust theory of harm: incentives to exclude downstream rivals Before intervening, a competition authority should try to understand why the vertically integrated firm would want to engage in a certain type of conduct. This consists of understanding whether there are legitimate reasons to adopt that conduct, or whether there exist anti-competitive reasons, that is, whether there is a clear theory of harm underpinning the behaviour of the vertically integrated incumbent.

For example, in an outright refusal to supply case, where a firm explicitly denies the input to a downstream rival, it would be difficult to argue that there was no incentive for that firm to exclude the downstream rival (otherwise it would have served it). Nevertheless, it is important to understand *why* there is refusal to supply. It may be that the buyer and the seller are not agreeing on the price of the input (as seemed to be the case in *Aspen*, as we shall see in Section 5.5, for example), and in that case one should be careful not to allow a party to seek antitrust intervention in relation to what is essentially a standard commercial negotiation or disagreement. For instance, it may well be that the price of an input is ‘high’ because the incumbent is (legitimately) trying to have the downstream buyers of its input to pay more for it, rather than trying to exclude them (whereas the latter may well argue the opposite, with a goal to minimise their cost base and have access to a key input on favourable terms).

As we discussed in Section 5.2 (and more formally in Section 5.3) the existing theories suggest that an incentive to engage in anti-competitive vertical foreclosure may arise due to either of two main reasons: the incumbent’s inability to extract a sufficient level of economic rents from the downstream rival or entrant; and the incumbent’s desire to protect or to transfer its market power in the future. Each of the different theories of harm belonging to either of these two broad categories relies on a number of specific conditions that would have to be met for the respective theory of

⁷⁵ The European Commission considered this in *Telefónica*, where it noted (para. 304 of the decision) that Telefónica’s incentives to invest in its infrastructure were not at stake, given that its infrastructure resulted to a large extent from investments undertaken before the introduction of broadband in Spain (that is, the service being investigated in that case) and that the original investments were made when Telefónica was benefitting from special or exclusive rights that shielded it from competition. The Commission also noted that under the investment criteria used by Telefónica at that time, the investments would have taken place even if there had been a duty to deal then.

harm to be plausible. As such, an authority would need to carefully check that the facts of the case are consistent with the conditions required by such theories of harm and that the mechanisms suggested by the economic models are plausible in the case at hand.

Under the first category of theories of harm, the incentive to vertically foreclose arises due to the inability by a vertically integrated firm to extract sufficient economic rents when considering to supply an input to (more or as efficient) downstream rivals: if the dominant firm is not able to appropriate a sufficient degree of the total profits available in the market, then it *may* find it more profitable to foreclose downstream rivals and supply the final market through its own downstream affiliate. Such inability to extract rents may be due, as we discussed, to upstream (wholesale) regulation, to the potential for opportunistic behaviour (giving rise to the ‘commitment problem’),⁷⁶ or to situations of uncertainty. So if in the assessment of an alleged case of vertical foreclosure one believed that this may arise due to imperfect rent extraction, then one should investigate whether this hypothesis is consistent with the facts.

Say, for example, that a vertically integrated firm not subject to regulation is already using contractual clauses that bypass the commitment problem. For the original hypothesis to hold, one would need to explain what makes it difficult for the vertically integrated firm to extract downstream rivals’ rents through suitable contracts.

Similarly, if one wished to put forward a partial foreclosure theory of harm based on an upstream rival or fringe, one would need to show that the vertically integrated incumbent was successful in committing not to deal with downstream rivals (even if it may have had the incentive to do so). For instance, in Section 5.2.3.4, we briefly mentioned how some economists (for example, Choi and Yi, 2000, and Church and Gandal, 2000) suggested that input design and compatibility choices may be one way for firms to commit not to deal with certain rivals.

Further, it is important to bear in mind that even the stylised models presented do not yield unambiguous conclusions across all circumstances. For example, in the discussion of how the existence of wholesale regulation may hinder the ability for an incumbent to extract downstream rents from

⁷⁶ Recall from Section 5.2.3.2 that the commitment problem arises where the vertically integrated incumbent is unable to commit to be inactive in the downstream market. Such problem limits the extent of the rent or payment that the incumbent is able to extract from downstream rivals, as these would anticipate that strong downstream competition would prevent them from earning sufficient profits to meet the payment requested by the incumbent in order to deal.

downstream rivals, we also noted how the incumbent would still have an incentive to serve the downstream rival (and extract an efficiency rent from it) should the rival be sufficiently more efficient than the incumbent's downstream affiliate. Moreover, the more differentiated the downstream firms, the less likely that vertical foreclosure would be profitable and thus that it would take place.⁷⁷

Another theory of harm we discussed in this chapter was instead based on a similar principle to that underpinning a number of theories of harm reviewed in the previous chapters in relation to other potential exclusionary practices: namely that of denying a rival the scale (or demand) it needs in order to be successful. As discussed in Section 5.2.4, and then shown more formally in technical Section 5.3.6, in a situation in which scale economies are important in one of the vertically related markets (say the downstream market), by denying the input to the rival currently, the incumbent prevents it from achieving efficient scale (or profits, or customer base) and makes it unprofitable for the downstream rival to operate also in the future. In turn, if a downstream rival's success is a pre-condition for an upstream rival's entry, by denying the input to the downstream rival currently the vertically integrated dominant firm may deter upstream entry in the future. (Or *vice versa*, if the success of an upstream rival is necessary for downstream entry, refusing to purchase from the upstream rival may be motivated by downstream entry deterrence.) Note that the theory is dynamic in nature: from a static point of view, the vertically integrated dominant firm would find it more profitable to supply the downstream rival, but this would entail forgoing profits in the future, when upstream entry would take place and would thus eliminate the upstream monopolistic position held by the incumbent. Foreclosure may therefore protect a monopoly position. In a slightly different version of the theory presented, we have also shown that if future upstream entry could not be deterred (that is, the upstream monopoly cannot be protected) then refusing the input to a downstream rival currently (which, again, would not

⁷⁷ The logic of imperfect rent extraction is very similar to the so-called 'vertical arithmetics' exercise performed in the context of vertical mergers to assess the incentive that the merged entity may have to foreclose downstream rivals. In essence, this amounts to assessing whether the potential losses from forgoing upstream (wholesale) revenues by not supplying a downstream competitor could offset or outweigh any increased downstream profits from the marginalisation or elimination of that downstream competitor. In a vertical merger, this assessment is prospective, that is, performed before the merger takes place.

be optimal in the short-run) may allow the vertically integrated dominant firm to ‘transfer’ or ‘create’ a downstream monopoly.

This (dynamic) theory of harm too is based on a precise mechanism, and for it to be convincing, one would have to verify that the facts of the case are consistent with such theory. First, scale economies need to be important in one of the vertically related markets and the rival active in that market (or seeking to enter that market) needs to achieve a critical scale so as to be successful. Second, future entry must be a material threat. For instance, if the dominant firm’s input were protected by a patent which was not going to expire over the following ten years and there was no obvious forthcoming substitute input, such a rationale for vertical foreclosure would be unlikely to apply (since upstream entry could not be expected within a reasonable time horizon).

Similarly, the idea that vertical foreclosure may protect an existing monopolistic position hinges upon the assumption that future upstream (respectively, downstream) entry depends on the success of an independent downstream (respectively, upstream) firm, something which should be checked against the facts of the case. Many factors may need to be considered for such an assessment, including the nature of the goods or services being sold, the nature of consumer demand (including possible network effects), barriers to entry and expansion and any technology or compatibility issues.

Furthermore for the dynamic theory to work – at least when scale economies arise from the supply side – the decision not to supply the independent downstream rival must be irreversible and have a commitment value. This is an important factor to assess.⁷⁸

Finally, one may also want to consider the ownership status of the potential entrants, as the incentive to engage in vertical foreclosure in a dynamic perspective is weaker when the potential entrants are vertically integrated.

Type of alleged foreclosure: refusal to supply versus margin squeeze In assessing whether the vertically integrated firm has an incentive to exclude a downstream rival, the difference between refusal to supply and margin squeeze should also play a role. As we highlighted in Section 5.2.3.1, by

⁷⁸ In the technical Section 5.3.6.3 we highlighted, though, that the commitment does not need to be eternal and thus not necessarily technological. Further, should network externalities exist in the downstream market, exclusion may be attained even without a commitment to refusal to supply. See the discussion in Section 5.2.4.

refusing to sell an indispensable input (or by setting a prohibitively high price for it), the vertically integrated firm makes it impossible for the rival to operate in the downstream market and, at the same time, it manages to set the monopoly price in that market.⁷⁹ If the vertically integrated firm cannot engage in refusal to supply, for instance because of the presence of a duty to deal (and cannot resort to more hidden practices that may achieve the same purpose, for instance quality or interconnection degradation), then in order to foreclose the activity of the downstream rival the vertically integrated firm must introduce a (downward) distortion to the retail price. Margin squeeze, then, is not a perfect substitute of refusal to supply. Indeed, Section 5.3.2.1 has shown that a margin squeeze would entail a profit sacrifice relative to the case of serving the rival. This observation does not imply that one will never observe a vertically integrated incumbent engaging in margin squeeze, but that the rationale for such a strategy would have to come from a *dynamic* context in which the profits sacrificed to squeeze the rival's margin can be recouped in later periods. Recoupment will be possible if the squeeze of the rival's margin in early periods leads to exclusion (or marginalisation) of the downstream rival, so that the vertically integrated firm can increase its retail prices in later periods.

These observations highlight a strong similarity with the conditions discussed in Chapter 1 for the rationality of predatory pricing. As a result, we expect to observe margin squeeze in situations in which the downstream rival needs to achieve a critical scale (or customer base) in order to be successful (or to expand), and the combination of retail and wholesale price chosen by the incumbent denies access to that scale. The existence of *significant economies of scale* in the downstream market – whether due to fixed costs, learning effects, demand externalities, two-sided market effects, or other reasons – is obviously key for this argument to apply. There must also be a *strong incumbency advantage*, which as a first approximation may be proxied by asymmetries in the investment in a crucial infrastructure and/or by persistent asymmetries in terms of market shares; this may be reinforced by the presence of significant switching costs, by the infrequency of purchases of the good or service sold and by demand externalities.

⁷⁹ Note that if a margin squeeze was mechanically defined as a situation where the combination of the vertically integrated incumbent's wholesale and retail prices fail the margin squeeze test (see the discussion in Section 5.4.3), constructive refusals to supply would also be caught within this definition (for example, when the rival is offered an extremely high wholesale price and rejects it). However, such a refusal to supply would not entail the profit sacrifice of a genuine margin squeeze for the incumbent and these two types of behaviour are therefore conceptually different.

Moreover, we expect margin squeeze to be more likely if the *downstream market is sufficiently mature*. This is because, first, a rapidly growing market is one where it is more likely that demand in later periods is sufficient for any downstream rival to reach efficient scale; second, in a growing market, the number of contestable buyers is likely to be larger relative to the captive ones, limiting the incumbency advantage. As a result, we believe that it is less likely that a vertically integrated firm has an incentive to engage in margin squeeze to exclude a downstream rival in industries where there are no entrenched market positions or where there is rapidly expanding demand. However, the same caveat we set out in Chapter 1 (in particular, in Section 1.4.2.2) for markets that undergo an evolutionary change (for example with the possible advent of a new technology) must be reiterated here.

Furthermore, we expect margin squeeze to be more likely if *financial frictions* are severe in the downstream industry and if there exist clear asymmetries between the incumbent's affiliate and the downstream rival in terms of ability to sustain losses and to raise external funds.⁸⁰

In sum, a coherent economic mechanism as well as facts consistent with this mechanism should be identified.

Upstream versus downstream dominance The above discussion suggests that an asymmetry between the downstream affiliate and the independent firm strengthens the incentive for the vertically integrated firm to engage in margin squeeze. Hence, a strong position of the incumbent's affiliate in the downstream market should make margin squeeze more likely to be profitable. However, this does not imply that downstream dominance should be regarded as a necessary prerequisite for intervention, as we suggest instead for predation and rebates (see Chapters 1 and 2).⁸¹ The incentive to engage in margin squeeze crucially relies on the existence of upstream market power that the vertically integrated incumbent is not able to fully exploit (due to imperfect rent extraction) or that it wants to protect

⁸⁰ For instance in the *Telefónica* case, that we discuss in Section 5.6.5, some downstream rivals were affiliates of international telecoms groups. One may argue that these firms may have received financial backing from the parent companies, thereby being in a stronger position to react to Telefónica's offers (and potentially sustain losses) than otherwise. There should be less scope for a margin squeeze in that environment, all else equal.

⁸¹ In the EU, a prerequisite for a finding of predation is that the incumbent's *downstream* affiliate is dominant, whereas for refusal to supply and margin squeeze cases only dominance in the *upstream* market is required. This might explain why some cases which appear similar were dealt with as different violations (see, for example, *Deutsche Telekom*, discussed later in this chapter, and *Wanadoo*, discussed in Chapter 1 on predation).

from future entry threats. Then, based on the economic models presented in this chapter, in margin squeeze and in refusal to supply cases it makes sense for antitrust intervention to require the existence of a dominant position in the upstream market but not necessarily in the downstream market. The extent of market power that the vertically integrated firm is able to exert in the downstream market is nevertheless an important element to take into account in the overall assessment of the incentives to exclude.⁸²

Type of intervention: antitrust versus regulation In the assessment of the incentives to engage in margin squeeze the focus is on the downstream market. Given the wholesale price that the vertically integrated firm sets for the input, the purpose is to assess whether this firm has an incentive to choose a retail price that makes it unprofitable for rivals to compete downstream. There is no *per se* contradiction, therefore, if a competition authority intervenes sanctioning a firm for margin squeeze in industries in which access to the upstream input is regulated by a sectoral authority.⁸³ For instance, in a market in which the wholesale price is subject to price cap regulation, the dominant firm has the ability to choose the retail price so as to exclude downstream competitors and still be compliant with sectoral regulation. An intervention of the antitrust authority against such a conduct would not express a conflict between the activity of the regulator and of the antitrust authority.

The logic is different if the choice of the wholesale price may suggest an exploitative abuse (see the discussion in Section 5.3.2). Under such a scenario, a vertically integrated firm would choose a 'high' wholesale price so as to extract rents from a more efficient downstream rival, not to exclude it. Hence, it is the sectoral regulator, if it exists, that should be in a better position to intervene as appropriate.

More generally, one may expect that the sectoral regulator has the right expertise and has the powers (for example, the ability to impose conditions on how certain firms deal with competitors and customers) to intervene to promote competition as appropriate and it is difficult to see how competition law should be expected to play a prominent role in markets characterised by such regulation. Granted, regulators may be less

⁸² In *TeliaSonera*, which we review in Section 5.5.2, the Court of Justice stated, however, that downstream dominance is generally not a relevant consideration in the assessment of margin squeeze.

⁸³ For an analysis of how to treat margin squeeze in a regulatory environment, see Bouckaert and Verboven (2004).

familiar in dealing with competition issues and may be more likely to be prone to capture (given that they deal with regulated firms on a regular basis), but on the other hand, regulators typically have a more profound knowledge of the sector, and may also have a longer-run perspective and may be more mindful of guaranteeing firms' incentives to invest, along with pursuing broader policy objectives. Hence, it would be difficult to justify, in general terms, the prevalence of a competition authority over a sectoral regulator in the context of the issues considered in this section.

Yet, there may be exceptions to this general principle. Imagine, for instance, that the regulated wholesale price is subject to retail-minus regulation, that is, it is computed by subtracting from the retail price an amount that ensures that downstream costs are covered. This type of regulation should avoid margin squeeze by its very nature and a competition authority should have no reason to intervene, unless it believes that the exact application of that type of regulation was incorrect and this may lead to anti-competitive harm through the dominant's firm conduct. This might have been the case in *Telefónica*: despite the fact that the wholesale price at the regional level was subject to retail-minus regulation, the European Commission intervened on the basis that the *ex ante* estimates of downstream costs used to compute the wholesale price turned out to be lower than the costs actually incurred. Hence, according to the Commission, there was scope for margin squeeze despite the existence of retail-minus regulation. There may also exist cases in which sectoral regulators take decisions which are clearly against competition principles, and we would find it difficult to exclude the possibility of intervening in such cases.⁸⁴ Such exceptions, however, are likely to be rare and any intervention by competition authorities that would cut across *ex ante* regulation would have to be very clearly motivated and reasoned through.

As we shall see in our review of the case-law (see Section 5.5), different jurisdictions have different rules with respect to the relationship between competition law and regulatory frameworks. For instance, in the EU there is a well-established principle that competition law prevails upon sector-specific regulation, and indeed in a number of cases (for instance in *Deutsche Telekom* and *Telefónica* that we shall discuss in Section 5.6) the dominant firms which have eventually been found guilty of abusive conduct had followed the prescriptions set by their respective national regulatory authorities.

⁸⁴ What if the regulator invites the regulated firms to jointly set prices? We would certainly not exclude antitrust intervention just because collusion was sponsored by a regulator!

Matters are different on the other side of the Atlantic. In *Trinko* (which we discuss in Section 5.5 in more detail), the US Supreme Court indicated that the role of the Telecommunications Act (1996) had to be taken into account before considering any antitrust intervention in the telecoms industry:

The [Telecommunications] Act's extensive provision for access makes it unnecessary to impose a judicial doctrine of forced access...Antitrust analysis must always be attuned to the particular structure and circumstances of the industry at issue. Part of that attention to economic context is an awareness of the significance of regulation...One factor of particular importance is the existence of a regulatory structure designed to deter and remedy anticompetitive harm. Where such a structure exists, the additional benefit to competition provided by antitrust enforcement will tend to be small, and it will be less plausible that the antitrust laws contemplate such additional scrutiny. (Parts III and IV of the Supreme Court's Opinion)

We shall return to the relationship between antitrust and regulatory intervention in Sections 5.5 and 5.6, in the context of specific cases that appeared before authorities or courts.

Remedies As already noted in Section 5.1, and in line with Werden (1987), we believe that competition authorities and courts need to consider the availability and viability of remedies to the behaviour they may try to curtail. In the context of refusal to deal, this would require considering, for example, whether mandating access may create additional costs to the incumbent or the rivals or entrants, create congestion or decrease product or service quality. An authority or court mandating access should also set out the exact terms for such access – imposing access does not imply that it should be given for free; rather, there should be a fair reward for access. Identifying such terms can be complex and an authority should be prepared to commit resources to monitor the remedies (or appoint a third party to do so).⁸⁵

⁸⁵ For example, in the *Microsoft* case (which we discuss in Section 5.6.1 in more detail), the European Commission found an abuse of dominance and imposed a remedy on Microsoft to provide certain interoperability information on reasonable and non-discriminatory terms. However, the Commission then came to the view that a very large part of the unpatented interoperability information provided by Microsoft as a result of that remedy lacked innovation; and that a comparison with the pricing of similar interoperability technology showed that royalties demanded by Microsoft for such unpatented information were unreasonably high. The Commission thus imposed a penalty on Microsoft, following further proceedings after the original Commission's investigation. See European Commission, MEMO/08/125, 27 February 2008.

Finally, in margin squeeze cases, if the squeeze is operated through low retail prices, the remedy would presumably consist of obliging the incumbent to increase its retail prices. One needs to be really sure about the margin squeeze infringement before imposing such a remedy that would result in consumers paying higher prices at least in the short run.

5.4.2.2 Efficiency and objective justifications

So far we have focused on the potential anti-competitive effects that may arise from practices falling within a vertical foreclosure framework. In the assessment of such cases (like with any other potentially anti-competitive practices, as acknowledged by authorities and courts across various jurisdictions) a competition authority should also examine any efficiency or objective justifications for the conduct at issue as put forward by the defendant, checking whether these flow from the conduct being investigated and whether they offset or outweigh any anti-competitive effects identified.

First of all, notice that there are a number of objective (or business) justifications that may explain a refusal to deal in a business context: some customers may be bad debtors; other customers may have failed to observe some contractual obligations; there may be a shortage of stocks; available capacity may be limited; the input may be in the process of being phased out; there may be safety or security concerns, or services may suffer from quality degradation, should access be granted.⁸⁶ Before starting a complex investigation into alleged vertical foreclosure, a competition authority should first verify that none of these legitimate justifications apply.

In certain contexts, reputation effects may also play an important role. To the extent that a downstream firm has a bearing on the quality of the final output or the experience of the final consumer (for example, a retailer's quality of service), a vertically integrated firm may be reluctant to deal with certain downstream rivals, which may hurt the integrated firm's reputation. Hence, reputation considerations may in principle justify refusal to supply. However, claims that an input is refused for reputation or safety reasons cannot be vague, and the dominant firm should support them with unambiguous evidence, such as a clear failure of the downstream firm to comply with the requisite standards.

As for cases of alleged margin squeeze, one may need to consider the possible role of a different type of efficiency defence. As discussed in Chapter 1 in relation to predatory pricing, low retail (downstream)

⁸⁶ See also Friederiszick and Gratz (2012) and O'Donoghue and Padilla (2013: 564–66).

prices may enable faster take-up of new products in nascent markets (this practice is sometimes referred to as ‘penetration pricing’) or more generally may enhance consumer goodwill, or may be part of the normal competitive process in markets characterised by features such as switching costs, learning economies, network effects, or two-sided externalities. This may benefit consumers in the long-run, as opposed to a counterfactual scenario where higher retail prices prevailed during the initial phase of a new market, and these slowed down the development of such market or potentially choked it off altogether.

In relation to the assessment of the legitimacy of low retail prices, it is worth considering one important difference relative to predation cases. As discussed above, in margin squeeze cases the vertically integrated firm does not have to be dominant in the downstream market. Hence, ‘low’ retail prices (such as penetration prices) may be a more valid justification than in predation cases where a firm is already dominant in the market where the alleged predation takes place.

5.4.3 Summary: Analysing Vertical Foreclosure Cases in Practice

Next, drawing from the discussion of the key factors for the assessment of alleged vertical foreclosure earlier in this section, we summarise the main steps that a competition authority may wish to take when investigating allegations of vertical foreclosure in practice. We start with refusal to supply, and then move to margin squeeze. As a premise, recall that antitrust intervention in this area should be restricted to cases in which the input is indispensable and is not the fruit of recent investment and innovation by the dominant firm (see Section 5.4.1), in order to preserve incentives to invest throughout the economy. Further, from the very beginning of any investigation, it is important to have a clear theory of harm, namely to understand for what reason the vertically integrated owner of an input may want to exclude possible downstream rivals, given that these are buyers of the input, and as such a source of profits. Finally, it would be desirable for a competition authority to discuss the case at hand with any relevant regulatory bodies, so as to avoid the risk of unwarranted inconsistencies in intervention, as well as to make sure that the competition authority is really the best placed body to intervene.

Analysing allegations of refusal to supply in practice In a case involving refusal to supply allegations, a competition authority may want to proceed as follows:

1. Determine any existence of a dominant position in the upstream market combined with indispensability of the upstream input for the possibility to viably compete downstream.
2. Identify any existence of a theory of harm explaining the reasons why the vertically integrated firm has an incentive to exclude downstream rivals. This requires to investigate whether the facts of the case match the key conditions of the theories that would rationalise the incentive to exclude.
3. Carefully evaluate possible objective justifications that may explain the business rationale of the practice at hand, and may offset or outweigh anti-competitive effects, along with an assessment of the need to preserve the incentives to invest/innovate.

Further, as noted above, it will be important for the authority to evaluate the availability, viability and proportionality of any remedies that may address any competition concerns identified.

Analysing allegations of margin squeeze in practice In a case involving margin squeeze allegations, a competition authority may want to proceed as follows:

1. Determine any existence of a dominant position in the upstream market combined with indispensability of the upstream input for the possibility to viably compete downstream.
2. Identify any existence of a theory of harm explaining the reasons why the vertically integrated firm has an incentive to exclude downstream rivals. Alternatively, the alleged margin squeeze may simply be the result of the vertically integrated firm seeking to make the buyers pay more for its input. As such, though, it would not be a possible case of exclusionary conduct but it may constitute an exploitative abuse (which may or not be infringing competition laws depending on the jurisdiction).
3. Check whether the margin obtained by the vertically integrated firm on the sales to the final market is lower than the margin obtained by supplying the input to the downstream rival, which represents the opportunity cost of downstream sales. In its simplest form,⁸⁷ this test is equivalent to checking whether $p - c < w$ where p is the retail

⁸⁷ See Bouckaert and Verboven (2004) for a more complete version of a margin squeeze test.

price, c is the downstream cost and w is the wholesale price.⁸⁸ If the above condition is satisfied, the vertically integrated firm sacrifices profits in the short-run in the attempt to foreclose the downstream rival.⁸⁹ Alternatively, this test can be interpreted as a replicability test: if $p - w < c$, a downstream competitor that is as efficient as the incumbent's affiliate and that acquires the input from the incumbent at the price w would not be able to match the incumbent's retail price and cover the downstream cost.

Given the similarity with predation cases, for a discussion on how to apply this test, we refer the reader to the discussion in Chapter 1. In particular, it may be appropriate to consider two benchmarks for the downstream cost. The upper bound may consist of average total costs ('ATC'), or long-run average incremental costs ('LRAIC'). The lower bound may consist of average variable costs ('AVC'), average avoidable costs ('AAC') or short-run average incremental costs ('SRAIC').^{90,91}

Then, if the difference between the upstream and the downstream price of the vertically integrated incumbent (that is, $p - w$) is larger than the upper bound of such cost range, the allegation should be

⁸⁸ In practice, firms typically produce several goods or offer several services, so one needs to determine the correct level of aggregation at which to perform a margin squeeze test. This problem is also related to the definition of a relevant antitrust market (for example in a predatory case). The aggregation level should be sufficiently high to constitute a relevant antitrust market. See, for example, Bouckaert and Verboven (2004). We discuss aggregation level issues in the context of *Deutsche Telekom* and *Telefónica* in Section 5.6.

⁸⁹ This test refers to a situation where final products are homogeneous and where the downstream rival and the incumbent's downstream affiliate use the same amount of input per unit of final product. If downstream products are differentiated and/or the downstream rival uses a lower amount of input than the incumbent's downstream affiliate, the opportunity cost of downstream sales decreases and the test should be adjusted accordingly. In order to account for this, Jullien et al. (2014) propose a test for margin squeeze modified on the ground of the diversion ratio (that is, the fraction of sales lost by the own affiliate that goes to the downstream rival) and of the input requirement of the downstream rival. However, these parameters reflect characteristics of the independent competitor that are difficult to be assessed by the integrated firm when it chooses retail and wholesale prices. This raises an issue of legal certainty and casts doubts on the implementability of the modified test.

⁹⁰ See Chapter 1 (in particular, Section 1.4.3) for a more detailed discussion on the differences between these notions of costs.

⁹¹ In practice, as also confirmed in our review of cases in Sections 5.5 and 5.6, authorities and courts (at least in the EU) have typically adopted LRAIC when implementing margin squeeze tests. This may be a result of many margin squeeze investigations hailing from the telecoms sector, where sector regulators had largely focused on this cost measure. However, in principle, there is no reason to rule out the consideration of more short-term cost measures, depending on the circumstances of a case. See also O'Donoghue and Padilla (2013).

dismissed. If the difference $p - w$ is smaller than the lower bound of the cost range, then there should be a presumption of margin squeeze, with the defendant having the burden of proving otherwise. Finally, if the difference $p - w$ falls between those bounds, for a margin squeeze finding, we think that it would be advisable from an economic perspective if the competition authority: (i) identifies the precise mechanism that would explain the ability and the incentive of the vertically integrated firm to exclude the downstream rival(s); and (ii) tests whether the broader evidence base supports the hypothesis that the vertical integrated firm had a coherent plan to sacrifice profits in order to exclude downstream rivals. An important part of this exercise is the assessment of whether the characteristics of the downstream market (including any scale economies, asymmetries between incumbent's affiliate and rivals) support the incentive to exclude.

It is important to recall again that the combination of the retail and wholesale prices may fail a margin squeeze test also in situations in which the vertically integrated firm is extracting rents from the independent (and more efficient) downstream competitor, with no exclusionary intent (see the discussion in Section 5.3.2.1). When this is the case, banning price combinations that fail the test set out above may introduce other distortions and may lead to ambiguous effects on total welfare. As a reaction to the ban, the vertically integrated firm may decrease the wholesale price, but it may also increase the retail price which will induce the competitor to do the same to the detriment of final consumers and welfare.⁹² This consideration emphasises the importance of not implementing the test in a mechanic way and of identifying a coherent and solid theory of harm.

4. Assess the degree of downstream dominance of the vertically integrated firm. Under a theory of harm where margin squeeze is used in a predatory fashion to exclude or marginalise a rival, the vertically integrated incumbent must have a degree of downstream market power both to constrain the downstream rival's prices and to recoup any losses once the rival has been excluded or marginalised.
5. Assess whether there may be any legitimate reason for the dominant firm to set prices below cost. As discussed above, for instance, if

⁹² The balance between these two effects depends on technology and demand. Jullien et al. (2014) provide examples in which a ban on margin squeeze may lead to higher or lower retail prices.

the firm is not dominant downstream, under certain circumstances penetration pricing may be justifiable.

We next turn to a review of the landmark case-law from both sides of the Atlantic, followed by a more in-depth review of selected cases, where we comment on whether the facts of those cases may have been consistent with some of the economic principles and theories of harm we discussed in this chapter.

5.5 Case-law

In this section we briefly review, in turn, the landmark cases on vertical foreclosure from the US and the EU. While the nature of antitrust intervention was broadly similar across the two jurisdictions at an early stage of enforcement, courts on either side of the Atlantic began drifting apart in their approach to enforcement over time.

As will become clearer, the four main differences between the US and the EU in their antitrust enforcement on vertical foreclosure issues can be summarised as follows. First, EU enforcers have developed a more interventionist approach than their US counterparts. Second, arguments supporting the supremacy of IP law over competition law (that is, protecting intellectually property rights regardless of the effects on competition in a market) have received more attention in the US than in the EU. Third, there is a clear supremacy of competition law over *ex ante* regulatory intervention in the EU, while this is not so in the US. Fourth, margin (or price) squeeze is a stand-alone anti-competitive conduct in the EU while not in the US.

5.5.1 United States

In this section we briefly describe how the approach of US courts in relation to vertical foreclosure has evolved over time. We begin with a review of important cases on refusal to deal (or to license), followed by a review of margin squeeze cases.⁹³

Refusal to deal/license cases The Supreme Court faced a refusal to deal case already back in 1912, in *Terminal Railroad*. There, a group of railroad companies (essentially acting as a single firm, referred to as the ‘terminal

⁹³ In this section we draw from a number of sources, including Areeda (1990), Carlton (2001), Gavil et al. (2008), Heimler (2010), Meese (2005), OECD (2009), Pitofsky et al. (2002), Salop (2010), Shelanski (2009), Waller and Tasch (2010) and Werden (1987).

company’) controlled railway bridges and switching yards leading to St. Louis, Missouri. The terminal company prevented competitors from using such infrastructure, keeping them out of the routes into and out of St. Louis. The Supreme Court found this behaviour to be an illegal restraint of trade and an attempt to monopolise the market and mandated access to the above-mentioned facilities on reasonable and non-discriminatory terms.

In 1973, in *Otter Tail*, the Supreme Court found that a vertically integrated electricity company (Otter Tail Power Co.) infringed the law by refusing to sell wholesale electricity and by disallowing the transfer of electricity from other sources over its own infrastructure to municipalities that wished to be active at the retail level.⁹⁴ Otter Tail was active in 465 towns at the retail level across Minnesota, North Dakota and South Dakota. It had enjoyed a legal monopoly (municipally granted franchises) in these towns for periods between 10 and 20 years. The refusal to deal took place at the end of such periods, when the markets were liberalised. In addition to the refusal to provide its own electricity, Otter Tail was found to have illegally denied access to its transmission network, thus preventing municipalities from purchasing wholesale electricity from neighbouring power companies. Similarly to *Standard Oil* (see our discussion in Chapter 1), Otter Tail was also found to have thwarted potential competition by taking over some of the municipal systems.

A few years later, an Appeals Court applied the essential facilities doctrine⁹⁵ in *MCI*, determining that it was illegal for a local telecommunications monopolist (AT&T) to deny interconnection to its own local network to competing long-distance carriers. As set out in Section 5.2.1, the Court reiterated the set of cumulative conditions that are necessary for a refusal to deal to infringe competition law: a monopolist controls the essential facility; the competitor cannot duplicate the essential facility; there is a denial of the use of the facility by the competitor; and it is feasible to provide access to it.

Two years after *MCI*, in 1985, the Supreme Court handed down a very significant judgment, *Aspen*. Aspen Skiing Company (‘Ski Co.’) had been offering joint marketing of a multi-day all-Aspen skiing pass, covering four areas together with Aspen Highlands Skiing Corp. (‘Highlands’). But by

⁹⁴ Similar facts occurred in *Mishawaka*. Differently from Otter Tail, however, the (vertically integrated) incumbent operator was found to have engaged in price squeeze, to the detriment of its downstream competitor (the municipal company of Mishawaka, Indiana).

⁹⁵ See Section 5.2.1.

1978 the parties could no longer reach a mutually satisfactory agreement on revenue-sharing. Highlands (operating one of the slopes) then proposed to purchase Ski Co.'s tickets at retail prices, in order to then offer skiers a bundled product with its own tickets. Ski Co. rejected this offer.

The Supreme Court found this anti-competitive (it found liability under Section 2 of the Sherman Act). It stated that a monopolist does not have a general duty to deal; however Ski Co.'s termination of an existing and presumably profitable course of dealing, together with its refusal to trade even at retail prices, was found to be exclusionary. The Court further explained that Ski Co.'s behaviour entailed actually foregoing sales; that Ski Co. had no efficiency justifications; and that the only advantage to it came from harming the competitor, which was found to have steadily lost market share.

Areeda (1990) underscored the insistence by the Court on the fact that Ski Co. had previously been dealing with Highlands, suggesting that dealing could be in the interest of both parties. This may in turn lead to the interpretation that refusals to *continue* to deal would attract antitrust scrutiny. Areeda (1990) argued, somewhat colourfully:

It may not have been wise for the Court [to stress the issue of discontinuance], because one of the consequences will be that lawyers will advise their clients not to cooperate with a rival; once you start, the Sherman Act may be read as an antdivorce statute.

Carlton (2001) raised some further insightful points in relation to *Aspen*. In his view, this matter was essentially a private dispute between two parties who were each seeking to obtain better terms of agreement. Antitrust should have played no role and Highlands should have been expected to lower its demands and accept an offer from Ski Co., since the next-best alternative (not dealing) was probably less profitable – unless (as it turned out) not dealing and seeking redress through an antitrust action was actually a more profitable option for Highlands.

More generally, Carlton (2001) argued (in a Coasian fashion) that there was no benefit from antitrust intervention since – absent antitrust intervention - the parties would have reached an agreement for the joint venture to continue if it was efficient to do so. He also lamented that the Court did not assess a different type of conduct, namely whether the rebates offered by Ski Co. to skiers in the form of discounted multi-day passes could in principle generate *de facto* exclusivity and foreclose Highlands in an anti-competitive fashion.

The door opened by *Aspen* to plaintiffs brought about a series of monopolisation cases in the following 15 years. The most notable ones related to the alleged refusals to license intellectual property rights.⁹⁶ If refusals to deal already caused controversy in US antitrust enforcement, the intellectual property dimension added a further layer of criticism against proactive antitrust enforcement: would a duty to deal not stifle the incentives to innovate in the first place? How could competition law enforcement sanction the very monopolies that IP law sought to create?

The cases that thus arose in the 1990s, as noted by Shelanski (2009), also clearly highlighted a systematic divergence between Appeals Circuits: while certain Circuits (for example, First, Ninth and Tenth) were open to a rule-of-reason assessment of refusals to license IP – that is, such refusals may, in principle, breach antitrust rules - the Federal Circuit essentially exempted such refusals to deal from antitrust scrutiny, suggesting a superiority of IP law over competition law.⁹⁷

Therefore, while in *Feist* and *Data General* the Appeals Courts did not find liability, these cases are nevertheless considered important because the judges at least assessed the potentially anti-competitive effects from a refusal to deal. *Feist* concerned a firm's refusal to sell its copyrighted directory listings to a competitor. The Court of Appeals (Tenth Circuit) reversed a lower court's ruling, mostly on the basis that the plaintiff had shown no evidence of anti-competitive effects flowing from the refusal to deal.

Two years later, in *Data General*, the Court of Appeals (First Circuit) considered a related matter. Grumman had complained that Data General, *inter alia*, had foreclosed it from the aftermarket for services offered to owners of Data General computers. The Court explicitly considered the (pro-competitive) business justification that a defendant may put forward when it holds IP: in the case at hand, an author's desire to prevent others from using its copyrighted work was a presumptively valid – but rebuttable – business justification, according to the Court. It found that the refusal to deal did not breach the monopolisation provisions. Shelanski (2009) inferred from this judgment that the court was implicitly suggesting the following logic: the more innovation is required to produce a good

⁹⁶ See Shelanski (2009) for a more detailed discussion.

⁹⁷ As noted on its website (www.cafc.uscourts.gov/), '[t]he Federal Circuit is unique among the thirteen Circuit Courts of Appeals. It has nationwide jurisdiction in a variety of subject areas, including international trade, government contracts, patents, trademarks, certain money claims against the United States government, federal personnel, veterans' benefits, and public safety officers' benefits claims.'

or a service, the greater the risk that a finding that a refusal to deal is anti-competitive will deter (pro-competitive) economic activity, thus the stronger the rebuttable presumption that an objective justification to refusing to grant access to IP is valid.⁹⁸

In *Kodak*, independent service organisations complained that Kodak unlawfully foreclosed them from the market for repair services of Kodak's photocopying machines.⁹⁹ Foreclosure allegedly arose because Kodak refused to provide spare parts of its own photocopiers to independent service organisations, thus preventing them from servicing Kodak photocopiers. Kodak justified its behaviour by asserting its IP rights. The Court of Appeals (Ninth Circuit) ruled against Kodak. As Shelanski (2009) wrote, the Court argued that:

[T]he protection of IP rights could be a valid business justification for refusing to supply a competitor but that the justification could fail if found to be a pretext for anticompetitive conduct.

In essence, the Court condemned Kodak's intent to use its monopoly in one market (that is, according to the Court, the market for Kodak photocopy and micrographics equipment parts) to create a monopoly in a second market (that is, the market for servicing Kodak photocopiers). The Court imposed a 10-year injunction forcing Kodak to deal.

The Federal Circuit took a very different stance on refusal to license. In *Intergraph*, Intel was appealing a lower court's preliminary injunction decision based on a finding of 'substantial likelihood' that Intel's behaviour (a refusal to supply to Intergraph) would constitute an infringement of competition law. The lower court observed that Intel, *inter alia*, had refused to provide Intergraph (an original equipment manufacturer) with special product specifications, which Intergraph deemed essential in order to compete in the graphics workstation business. Intel was deemed to have withdrawn its technical assistance and special benefits, including pre-release access to Intel's new products. The Appeals Court (Federal Circuit) disagreed with the lower court. One of its key arguments was that Intel and Intergraph were not competing in the same market, thus there

⁹⁸ This type of efficiency defence is what Areeda (1990) called a 'macro level' legitimate business justification (as opposed to 'micro level' business justifications from refusing to deal, which the author suggested could include more context-specific reasons, such as the failure to reach satisfactory terms of business, or the risk of causing harm to end-consumers).

⁹⁹ See Chapter 4 for a discussion of earlier proceedings before the Supreme Court with respect to tying.

was no sense in which Intel could ‘foreclose’ Intergraph by denying access to an ‘essential facility’.

In 2004 the Supreme Court appeared to depart from the ‘spirit’ of *Aspen* and handed down the *Trinko* judgment, which represents the current standard on refusal to deal in the US.

The case was about the decision by Verizon (an incumbent telecoms operator) not to offer full access to its local loop to emerging local competitors; more specifically, it was found to be providing them with a downgraded service. The question was whether this constituted a breach of the monopolisation provisions (irrespective of any regulatory duty that Verizon might have had under the Telecommunications Act 1996). The Supreme Court ruled that it did not. In doing so, however, it was careful not to directly contradict *Aspen*. Rather, as Fox (2005) noted, it sent a clear message to potential plaintiffs and lower courts that only complaints meeting the (very narrow) *Aspen* conditions could be successful, at least in principle. That is, monopolists have a broad freedom not to deal, except in very narrow circumstances.

The conditions that the Supreme Court used to set *Trinko* apart from *Aspen* were essentially three: first, in *Aspen*, the defendant had discontinued the supply of a service to a competitor, as opposed to a *de novo* refusal to deal (the Court would supposedly interpret discontinuation of supply as ‘evidence’ that dealing with the competitor must have been profitable at some stage, so that termination to deal could be evidence of an exclusionary motive). Second, *Aspen Skiing* had refused to offer a competitor access to its slopes even at the retail price, thereby making a profit sacrifice. Third, in *Trinko*, the Supreme Court indicated that the role of the Telecommunications Act (1996) had to be taken into account and noted that any additional benefit to competition from antitrust enforcement would be small.

Margin squeeze cases We are aware of few key US cases that focused specifically on margin squeeze allegations (that is, separately from refusal to deal or predation).

In *Corn Products* (dating back to 1916), a District Court found that the Corn Products Refining Company engaged into practices that were in breach of both restraints of trade and monopolisation provisions, including price (or margin) squeeze, by leveraging its monopoly on glucose and its high market share in starch to eliminate downstream competitors.

Later, in the mid-1940s, Alcoa was found to have infringed competition law through a margin squeeze in *Alcoa*. To our knowledge, the approach

by the Court of Appeals represented the earliest application of a ‘margin squeeze test’, which is the current practice in the EU (see Section 5.5.2). Alcoa was a monopolist in the market for aluminium ingot, which it also used to produce (aluminium) sheet. At the same time it supplied ingot to rivals in the aluminium sheet market. Rivals alleged a margin squeeze.

The Court computed the sum of the wholesale ingot price and Alcoa’s rolling (that is, ‘downstream’) costs and compared this with Alcoa’s sheet prices, across various gauges and for several years. It concluded that the margin was either negative or just about positive, thus ruling against Alcoa.

Over 60 years after *Alcoa*, the Supreme Court took a completely different stance in *linkLine*. This case related to an allegation of price (or margin) squeeze in the broadband market in California, where AT&T was the only wholesale provider and was also active at the retail level.¹⁰⁰

The lower courts had found in favour of the (four) plaintiffs, who were alleging price squeeze by the incumbent. The Supreme Court reversed. In doing so, it set a clear (and very narrow) test that had to be passed for a margin squeeze claim to be (potentially) successful.

First, the incumbent’s retail prices had to be predatory, according to the *Brooke* standards: prices had to be below-cost and there had to be a ‘dangerous probability’ that the incumbent could recoup losses after the competitor exited the market (see our discussion on predation in Chapter 1 for further details).

Second, the plaintiff had to successfully argue that the incumbent had an antitrust duty to deal at the wholesale level: absent that, there can be no margin squeeze, since the incumbent could always (lawfully) deny the input to the downstream competitor altogether.

Even if this test was passed, the Court continued, it would be problematic to assess what a fair margin for the competitor would be. But if it could be shown that both the wholesale and the retail prices were independently lawful, then even a wholesale price in excess of the retail price could not be evidence that the monopolisation provisions had been breached. This logic therefore ruled out margin squeeze as a stand-alone anti-competitive practice in the US. This stands in contrast with the approach in the EU, as we discuss next.

¹⁰⁰ Although not key to this antitrust case, it is interesting to note that, following an earlier merger commitment, AT&T had undertaken not to price its retail services below its wholesale services.

5.5.2 European Union

In this section we provide a historical account of the approach followed by the European Commission and the Courts in relation to alleged vertical foreclosure through unilateral conduct. We begin by considering refusals to deal, continue by making some brief comments on standard-essential patents cases and then turn to margin squeeze cases, before concluding with a brief discussion of other forms of vertical foreclosure which the European Commission has scrutinised in the energy sector.¹⁰¹

Refusal to deal/license cases As Korah (2007: 172), among others, noted, the decisions and judgments by the Commission and Courts in early cases were mostly based on protecting downstream competitors. This approach changed with time, shifting the focus to the protection of the competitive process.

The first European landmark case in this area of competition law is probably *Commercial Solvents*. According to the Commission, Commercial Solvents was the only producer in the world of the raw materials needed to produce ethambutol (used to make certain medicines). It was also vertically integrated, in that it controlled Istituto Chemioterapico, which manufactured ethambutol and supplied it across the (European) Common Market at the time. The Commission found that Commercial Solvents abused its dominant position in the market for raw materials necessary to produce ethambutol by refusing to supply Zoja, one of the only two other main manufacturers of ethambutol in the Common Market. On appeal, the Court of Justice upheld.

As we noted in Section 5.2.1, the Commission developed its essential facilities doctrine in the 1990s. To our knowledge, the first instance where the Commission referred to this notion was *Holyhead*, a case relating to ferry services from Wales to Ireland. Sealink owned the port infrastructure and also operated ferry services to Ireland. B&I provided competing ferry services and complained that Sealink scheduled its own services to generate disruption to B&I's loading and unloading operations. This was not a classic case of refusal to supply, but in any event, it was the first instance in which the Commission applied the essential facility doctrine to justify antitrust intervention. In *Port of Rødby* the setup was similar. In this case, a State-owned company was the vertically integrated operator, that is, both owning the port and providing ferry services. The Commission found that

¹⁰¹ In this section we also draw from Ezrachi (2016), Geradin et al. (2012), Korah (2007), O'Donoghue and Padilla (2013), Russo et al. (2010) and Whish and Bailey (2015).

the refusal by the Danish Government to authorise the provision of access to such port facilities to a competing ferry operator (that is, a downstream competitor) constituted an abuse of a dominant position. In *Frankfurt Airport*, the Commission found that the owner and operator of the airport had a duty to grant access to ground handling services to competing firms wishing to operate there.

However, arguably, some of the most significant (and consequential) European cases on refusal to supply arose beyond traditional 'physical' essential facilities, as we discuss next.

Magill is a milestone in the EU case-law on refusals to supply. In the mid-1980s, there were only three TV stations in Ireland and each published its listings separately on a daily basis. Magill saw a business opportunity and started producing a weekly programme guide covering the three channels. The three broadcasters sued Magill, separately, for breach of copyright law. Magill retorted that this refusal to deal amounted to an abuse of a dominant position (where the dominance allegedly stemmed from each of the broadcasters having a *de facto* monopoly over the production and first publication of its weekly TV listings) and the Commission agreed with it. Both the General Court and the Court of Justice upheld on appeal. The Courts referred to three conditions that would lead to a refusal to supply by a dominant firm being abusive. First, there had to be no objective justification.¹⁰² Second, all competition on the ancillary market would have been eliminated following the refusal to supply an indispensable input. Third, the Court of Justice introduced a novel element with respect to the existing case law: the 'new product' test. In other words, the practice was found to be abusive because it prevented a downstream firm from offering a new product (combined weekly listings) for which there was considerable demand, and which would thus benefit consumers.

Bronner is another landmark case, arisen as preliminary ruling based on a reference from an Austrian Court. The Court of Justice took a different stance from that in *Magill*. The Advocate General's Opinion (issued prior to the Court of Justice ruling) made a clear distinction between favouring the competitive process and supporting competitors. It also stressed the importance of leaving intact the incentives to invest: mandating access to existing facilities, where building alternative ones is viable, lowers such

¹⁰² The Court of Justice did not set out in general terms what would constitute an objective justification. In the case at hand, it referred to the notion that the refusal to deal had no justification either in the activity of television broadcasting or in that of publishing television magazines.

incentives. The Austrian case arose as a complaint by Oscar Bronner GmbH & Co. KG, which sought one of its newspapers (Der Standard) to be home-delivered using Mediaprint's network (Mediaprint was also in the newspaper business with its own dailies, Neue Kronen Zeitung and Kurier). The Court of Justice agreed with its Advocate General and further clarified that a dominant firm has a duty to deal only if it has an essential facility (and if there are no objective justifications for the refusal). By contrast, in the case under review, it appeared that Der Standard could still compete in the newspaper market through alternative distribution networks (for example, through kiosks or by mail).

In sum, the conditions for a refusal to supply to be abusive, as laid out in *Bronner* were: indispensability of the input,¹⁰³ elimination of all competition and absence of an objective justification.

IMS Health also arose as a reference to the Court of Justice, this time from a German court. Based on the information reported in the judgment, the facts appeared to be as follows. IMS and NDC were both in the business of tracking pharmaceutical sales. IMS had developed its own 'brick structure' of reporting sales to pharmaceutical companies, that is, it had divided Germany in 1860 geographic areas (that is, the 'bricks') and it classified sales according to that grid. IMS also shared this brick structure with pharmacists and doctors, who in turn helped IMS improve its classification system. This way of organising data was very successful and the IMS brick structure soon became the industry standard in Germany. When NDC entered the market, it first used a different way of classifying data, but ended up adopting IMS's brick structure as customers appeared to be used to it. IMS sought an injunction from a local court to force NDC

¹⁰³ 'Indispensability' of the input was defined as follows: 'For [access to the dominant firm's distribution network] to be capable of being regarded as indispensable, it would be necessary at the very least to establish [...] that it is not economically viable to create a second home-delivery scheme for the distribution of daily newspapers with a circulation comparable to that of the daily newspapers distributed by the existing scheme' (para. 46 of the Court's judgment). The Advocate General, in his Opinion, had further stated: 'It seems to me that [requiring a dominant undertaking to supply the product or service or allow access to the facility], whether understood as an application of the essential facilities doctrine or, more traditionally, as a response to a refusal to supply goods or services, can be justified in terms of competition policy only in cases in which the dominant undertaking has a genuine stranglehold on the related market. That might be the case for example where duplication of the facility is impossible or extremely difficult owing to physical, geographical or legal constraints or is highly undesirable for reasons of public policy. It is not sufficient that the undertaking's control over a facility should give it a competitive advantage' (para. 65 of the AG Opinion).

to discontinue such usage. The local court referred the case to the Court of Justice for a preliminary ruling.¹⁰⁴

In this case the Court of Justice took the opportunity to clarify that the three conditions set out in *Magill* were sufficient for a finding of an abuse, but were also cumulative (and left the German court to evaluate them on their actual merits based on the factual evidence of the case). In the *IMS* case, though, it was fairly apparent that NDC was not introducing a new product but was rather offering very similar services as IMS, failing as a consequence the ‘new product test’.¹⁰⁵

One of the highest-profile antitrust cases in the EU, *Microsoft*, referred to a refusal to deal.¹⁰⁶ The Commission found that Microsoft abused its dominant position in the market for work group servers.¹⁰⁷

In short, Microsoft was found to have abused its dominant position by refusing to provide key specification protocols to competitors, thus denying interoperability of their software with Microsoft’s operating system. In its judgment upholding the Commission’s Decision, the General Court set out the conditions under which a refusal to deal would fall foul of the EU abuse of dominance provisions. These conditions may be summarised as follows. First, the refused input is indispensable for downstream rivals to compete viably. Second, there is a risk that the refusal to deal eliminates all effective competition in the downstream market. Third, the refusal to deal limits technical development to the detriment of consumers (this appeared to represent an evolution of the ‘new product test’ referred to above). Fourth, there are no objective justifications for the refusal to deal. We discuss this case in detail in Section 5.6.1, offering a possible interpretation from an economic perspective.

Finally, a series of cases that have been attracting substantial interest from the media and practitioners alike are those involving Google. As this book was going to press, in *Google (comparison shopping)*, the

¹⁰⁴ NDC also complained to the European Commission, which granted interim measures ordering IMS to grant a licence to use its brick structure to all the undertakings present on the market for the provision of German regional pharmaceutical sales data. These were later dropped after an appeal.

¹⁰⁵ Still, it is not fully clear whether IMS’s ‘brick structure’ was really an innovation worthy of the strong copyright protection awarded by the courts. Rather, this may look like a case where network effects and switching costs may make it more difficult for potential entrants to break into the market.

¹⁰⁶ We discuss the tying of Windows Media Player, the other major leg of this case, in Chapter 4.

¹⁰⁷ These performed routine tasks for small and medium businesses, such as routing jobs to printers.

Commission had just decided that Google had abused its dominant position by systematically favouring its comparison shopping service over those of competitors in Google's search result pages. In Section 5.6.8 we offer a summary and a preliminary interpretation of this case (noting that the Commission's decision has not been published yet at the time of writing).¹⁰⁸

Abuses related to standard-essential patents We next review two recent high-profile cases on standard-essential patents ('SEPs'), and discuss the extent to which they may also be seen as falling within the framework explored in this chapter. According to the Commission, SEPs are patents that are essential to implement a specific industry standard, so that it would not be possible for a manufacturer to produce goods that comply with a certain standard without accessing these patents (for example, through a licence).

The Commission recognised that seeking injunctions before courts (that is, put very informally, asking a court to force a firm that does not have a licence to use a patent required to manufacture a certain product to stop manufacturing it) was generally a legitimate remedy for patent-holders, in the case of genuine patent infringements.

However, the Commission also formed a view that seeking an injunction based on SEPs may constitute an abuse of a dominant position (i) if an SEP-holder had given a voluntary commitment to license its SEPs on fair, reasonable and non-discriminatory ('FRAND') terms over the course of the process that led to the setting of that industry standard and (ii) where the company against which an injunction was sought was willing to enter into a licence agreement on such FRAND terms (including where the licensee had agreed that it would accept the FRAND terms set by a court following a dispute with the patent-holder over such terms).

The Commission reasoned that, in such circumstances, injunctions could exclude rival products from the market and this may be anti-competitive.

The Commission appears to have applied the above test in both *Samsung* and *Motorola Mobility*.¹⁰⁹

¹⁰⁸ We mention further ongoing investigations involving Google, at the time of writing, in Chapters 3 and 4.

¹⁰⁹ See, respectively, European Commission, Press Release IP/14/490 and Press Release IP/14/489, 29 April 2014; and the press conference by Commissioner Almunia on 29 April 2014 'Introductory remarks on Motorola and Samsung decisions on standard essential patents'. The Commission also found that it was anti-competitive for Motorola to insist, under the threat of the enforcement of an injunction, that Apple should give up

In *Samsung*, the Commission was concerned about Samsung's use of injunctions (against Apple) in relation to its SEPs on the European Telecommunications Standardisation Institute's ('ETSI') 3G UMTS standard, a key industry standard for mobile and wireless communications. To address the Commission's concerns, Samsung committed not to seek injunctions for five years in Europe on the basis of SEPs for smartphones and tablets against any potential licensee that agreed to accept a specific licensing framework (with a negotiation period of up to 12 months and, if such negotiation failed, a court or an arbitrator setting FRAND terms). The Commission agreed to close its case against Samsung.

In *Motorola Mobility*, the European Commission was concerned about Motorola Mobility's injunctions against Apple in relation to GPRS, a 2G mobile and wireless communications standard. The Commission, applying the test set out above, found that it was abusive for Motorola to first seek and then enforce an injunction against Apple on the basis of an SEP which Motorola had committed to license on FRAND terms and where Apple had agreed to take a licence and abide by a determination of the FRAND royalty rates by a relevant court.

These SEP-related cases are particular cases which are, in our view, difficult to classify according to the theories and practices we have discussed so far. In a sense, they may be interpreted as potentially 'exclusionary' practices. This is because vertically integrated companies like Samsung or Motorola (the latter now part of Google) own essential patents (upstream) and manufacture the mobile devices (downstream). So, in principle, a firm in a similar position may attempt to use injunctions in order to 'exclude' certain mobile devices produced by downstream rivals and gain market power downstream (for example, by relegating competitors to a fringe or delaying their product development or access to key markets). These practices may alternatively (or possibly additionally) be assessed as possible exploitative practices, with injunctions acting as a tool to extract rents from downstream rivals in excess of what a FRAND remuneration of the patent would yield. These practices may also entail reputation effects for the patent-holder, which may be interpreted through either an exclusionary and/or an exploitative lens. These policy questions are relatively novel and quite complex, and it appears to us that more time (as well as academic research) is needed to develop a coherent framework to assess them.

its rights to challenge the validity or infringement by Apple's mobile devices of Motorola's SEPs.

Margin squeeze cases We next consider margin squeeze cases. An early case where the Commission found this behaviour to be a stand-alone abuse was in 1988, in *Napier Brown/British Sugar*. According to the Commission's decision, in the 1980s British Sugar was the largest sugar manufacturer and seller in the UK, controlling 58 per cent of the relevant market defined by the Commission (granulated sugar). Napier Brown was instead a merchant specialised in purchasing bulk quantities of sugar and then distributing it. After Napier Brown entered the retail market in 1983, British Sugar reportedly began engaging in a number of business practices. The Commission, seemingly implementing a 'margin squeeze test' (although it did not give any name to the methodology adopted), found that British Sugar left a margin between its prices for retail and industrial sugar that was insufficient to cover the dominant firm's (downstream) distribution costs. We discuss this test further in the context of the review of *TeliaSonera* below.

Next, we turn to three key margin squeeze cases, all from the telecoms industry (somewhat interestingly, Veljanovski (2013) noted that over the period from 2003 to 2012, over two-thirds of the 41 European cases on margin squeeze he reviewed took place in the telecoms industry).¹¹⁰

In *Deutsche Telekom*, the Commission found a margin squeeze by the German incumbent in the market for the access to the local telephony network (the decision was confirmed by both the General Court and the Court of Justice). In *Telefónica*, the Commission found a margin squeeze in the Spanish broadband market. The General Court and the Court of Justice upheld. (We review both these cases in more detail in Section 5.6.)

Both incumbent telecoms operators were subject to fairly tight *ex ante* regulation, which added a further layer of complexity. However, the European Commission (later supported by the courts) clearly affirmed the superiority of competition law over *ex ante* regulation (that is, compliance with regulation does not guarantee immunity from competition law), in sharp contrast to the US approach in *Trinko* previously discussed.

In *TeliaSonera*, handed down in February 2011, the Court of Justice (upon request by a Swedish Court) had the opportunity to set out the conditions under which a margin squeeze would take place and a dominant firm would be abusing its position (note that in the EU a margin squeeze

¹¹⁰ See also Geradin and O'Donoghue (2005) for a discussion of several national margin squeeze cases in the telecommunications sector.

has been deemed to be in itself capable of constituting illegal behaviour, that is, it is a stand-alone abuse).¹¹¹

The Swedish Competition Authority had found that TeliaSonera, the incumbent, had set broadband internet wholesale and retail prices between 2000 and 2003 in such a way that it could not have covered its downstream costs; that is, it would have squeezed the margins of an as-efficient competitor.¹¹²

The Court of Justice noted the following. First, a margin squeeze occurs if the difference between the upstream (input or wholesale) price and the downstream (retail) price for a good or service is either negative or insufficient to cover the costs that are relevant for the dominant firm to supply that good or service to end-users. The Court therefore pointed out that the relevant costs for the assessment are those of the dominant firm itself. (Note that this increases the level of legal certainty, as it is typically much more straightforward for a business to make decisions based on its own costs, rather than on its estimation of competitors' costs.)¹¹³ Second, for an abuse of dominance finding, it is necessary to demonstrate that, taking particular account of whether the wholesale product is indispensable, the pricing practice at issue produces an actual or potential anti-competitive effect on the retail market. Third, a margin squeeze may generate anti-competitive effects, at least in principle, even if the input is not indispensable. This appears to us a major departure from the case-law on vertical foreclosure, in that input indispensability remained a necessary condition for a refusal to supply to be abusive, but not for a margin squeeze to be abusive. Fourth, for any finding of an abuse of dominance to be justified, any efficiencies from the pricing practice at issue which benefit consumers must not offset or outweigh the exclusionary effect of the practice.

¹¹¹ While the *Telefónica* judgments were given later, our understanding is that they did not alter the substance of the legal test set out by the Court of Justice in *TeliaSonera*, in fact they extensively referred to it. The *Telefónica* judgments, in particular the General Court's, provide a deeper assessment of the facts of the case without setting out a complete legal test (we consider this case in more detail in Section 5.6.5).

¹¹² Note, however, that TeliaSonera had no regulatory duty to offer wholesale ADSL products.

¹¹³ The Court added that the costs and prices of competitors may be relevant to the assessment of an alleged margin squeeze, for example if the cost structure of the dominant undertaking is not precisely identifiable (see para. 45 of the preliminary ruling, which considers possible exceptions to the use of the dominant firm's own cost structure for the assessment of the alleged margin squeeze).

Controversially, in our view, the Court also made the following point: if the dominant firm's downstream margin was shown to be able to cover the firm's downstream costs, anti-competitive effects could nevertheless arise, at least in principle, so long as the dominant firm's pricing behaviour made it more difficult for rivals to compete in the relevant markets (for example because they would achieve lower profitability than otherwise).¹¹⁴ In our view, intervention by an authority or a court in such a scenario would be inconsistent with an application of sound economic principles and would as such be a poor policy for the promotion of competition. In fact, such a policy may have the effect of chilling beneficial competition in a number of circumstances. That is because even competition on the merits (for example, the introduction of a superior product by the dominant firm, or fierce downstream price competition) would be likely to lower rivals' profitability, compared to a counterfactual scenario with softer downstream competition.¹¹⁵

European energy cases following the Commission's Energy Sector Inquiry

In 2005, the European Commission launched an inquiry into competition in gas and electricity markets, which lasted until January 2007.¹¹⁶ The Commission concluded that the markets were not working effectively and opened various proceedings. One of its major concerns related to alleged anti-competitive foreclosure arising from vertical integration by certain energy incumbents. Based on the Commission's assessment, a common theme across a number of cases was the limited capacity available to competing downstream operators, as a result of the downstream affiliates of the incumbents being signed into long-term contracts with the upstream (monopoly) incumbent (the issue was referred to as 'long-term capacity bookings').

¹¹⁴ See para. 74 of the Court's preliminary ruling.

¹¹⁵ The Court of Justice also set out a number of factors that as a general rule are *not* relevant to the assessment of whether an alleged margin squeeze may constitute an abuse: whether the defendant is dominant in the downstream (for example, retail) market; whether it can be shown that the dominant firm would be able to recoup, in the future, any losses incurred through the margin squeeze; the degree of upstream dominance by the vertically integrated firm; the absence of regulatory obligations upstream; whether a new technology with large investments required is involved; or whether the pricing practice affects new and/or existing customers of the dominant firm.

¹¹⁶ See European Commission, 'Competition: Commission Opens Sector Inquiry into Gas and Electricity', Press Release IP/05/716, 13 June 2005; and 'Competition: Commission energy sector inquiry confirms serious competition problems', Press Release IP/07/26, 10 January 2007.

Three notable cases were *GDF*, *E.ON* and *RWE*. In these cases, the Commission was also concerned about other practices, including anti-competitive rebates, margin squeeze, capacity degradation and strategic (upstream) underinvestment. These cases led to the parties offering major commitments, in the form of either behavioural or structural remedies (including vertical separation). These eventually addressed the Commission's competition concerns, so that it closed those cases without any infringement finding.^{117,118}

We will review such cases, among others, in more detail in Section 5.6.6. At this stage, we simply note that some commentators (see, for example, Broomhall et al., 2012) pointed out that some of the theories of harm put forward in these cases, such as 'strategic underinvestment', were novel in EU case-law. These commentators therefore questioned whether the European Commission adopted such a type of decision strategically, in that this allowed the Commission to (by and large) avoid judicial review, while intervening quite heavy-handedly in a market, putting forward novel theories of harm. Other commentators (see, for example, Von Rosenberg, 2009) noted that the European Commission had never gone as far as imposing structural remedies as part of an antitrust investigation (that is, excluding merger control), and found it noteworthy that the Commission accepted the upstream network divestments offered in *E.ON* and *RWE* in return for closing the investigations without finding an infringement.

5.6 Cases

In this section, we review in some detail a number of cases from various jurisdictions and covering a range of practices that may fall under the vertical foreclosure framework.

¹¹⁷ Under European Council Regulation 1/2003, Article 9, parties may offer commitments (sometimes also known as undertakings), and if these address the competition concerns by the European Commission, the Commission may close its investigation without an infringement finding. These therefore expose the relevant companies to lower reputational risk and shorter proceedings than what a full investigation would entail, as well as generally give lesser grounds for follow-on antitrust damages. We also understand that the degree of judicial review is typically lower in the case of a commitments decision than in an infringement decision, which may allow the Commission to be more audacious in cases resolved through commitments decisions.

¹¹⁸ We note that the Commission released *E.ON* from its commitments almost five years ahead of schedule, as it believed that as a result of their implementation competition in the market had increased significantly. See European Commission, Press Release IP/16/2646 of 26 July 2016.

5.6.1 Microsoft (EU – Interoperability)

One of the antitrust cases that received most attention in the 2000s is *Microsoft*. Here, we focus on the interoperability part of the case, where the European Commission found an abuse of dominance with regards to a refusal to disclose information required for interoperability by competing workgroup server operating systems ('OSs').¹¹⁹

Workgroup servers perform basic IT tasks in a network connecting clients: they allocate tasks, handle security, share files, route print jobs.

The Commission argued that without sufficient disclosure of the protocols (interface information) needed for both server-to-server and client-to-server interoperability, Microsoft's competitors could not develop effective products, as customers would need seamless interoperability between clients and servers.

Microsoft had a very high market share in client PC operating systems (over 90 per cent), and the Commission argued that it denied or reduced interoperability to leverage its market power from client OSs to workgroup servers.

The Commission also found further consumer harm in the form of less product differentiation (than in a competitive market), since customers remained by and large locked into Microsoft's homogeneous suite of products, as a result of its abusive behaviour.

The Commission fined Microsoft €497m (this included a fine for the illegal tying of Windows Media Player) and forced Microsoft to offer competitors its protocol specifications for interoperability on reasonable and non-discriminatory terms.¹²⁰

Microsoft appealed the decision. One of Microsoft's main grounds of appeal was that its competitors were not going to offer a 'new product'.

In 2007, however, the General Court upheld the Commission's Decision, in a judgment seen by some commentators¹²¹ as merely casting the facts of the case (as set out in the Commission's decision) within the frameworks established by *Magill* and *IMS*. The General Court countered Microsoft's point on the 'new product': its competitors were not seeking to create a clone, but products which would be functionally equivalent. Further, the

¹¹⁹ The Windows Media Player tying allegations of the European case against Microsoft are discussed in Chapter 4.

¹²⁰ In 2008 the Commission imposed a penalty on Microsoft of €899m, arguing that Microsoft had not complied up until late 2007 with the obligations set out in the Commission's decision in relation to server interoperability (the General Court lowered this amount on appeal to €860m).

¹²¹ See, for example, Ahlborn and Evans (2009) and Larouche (2008).

General Court confirmed the Commission's broad interpretation of the 'new product' test: preventing the emergence of a new product on the market should be considered in the context of the EU's abuse of dominance provisions, which prohibit practices that '[limit] production, markets or technical developments to the ... prejudice of consumers' (para. 643 of the judgment).

The General Court also agreed with the Commission's implementation of a two-stage approach in the determination of whether the information at issue was indispensable. According to the General Court, the Commission was correct in first assessing the degree of interoperability necessary for a competitor to remain viable in the market; and in then assessing whether the withheld information was the only economically viable source to achieve the degree of interoperability determined at the earlier stage.

Further, the General Court also espoused the Commission's arguments that (i) there was a risk that competition would be eliminated as a result of Microsoft's behaviour;¹²² and that (ii) the exceptional circumstances of the case made it such that Microsoft's justification based on exercising intellectual property rights failed to justify the contested conduct.

In a case of refusal to deal or, like here, refuse to offer (seamless) interoperability, the crucial questions are (i) whether there could be an anti-competitive rationale for such a practice, and (ii) whether such a practice can be justified by the legitimate protection of intellectual property rights.

Kühn and Van Reenen (2009) argue that the *Microsoft* (interoperability) decision could be interpreted in the light of the dynamic theory set out in Carlton and Waldman (2002), which we first discussed in Chapter 4 (in a tying setting) and further in this chapter (see Section 5.3.6). The main trait of the theory is that the dominant firm intends to exclude a rival in a vertically related or complementary market in order to avoid that a successful product in such a market may develop into a future threat in the market where it holds a dominant position.

According to Kühn and Van Reenen (2009), server OSs could have developed into an alternative platform to PC clients running on Windows, because developers could write their software to the open standards and

¹²² Once achieved a large customer base, network effects would have been self-reinforcing, according to the Commission: a larger customer base meant that technicians had a stronger incentive to learn how to manage the Microsoft work group server operating system; and the more technicians available, the stronger the incentive for customers to choose Microsoft's workgroup server operating system.

application programme interfaces typically used by server OSs, thereby bypassing Windows. In this perspective, Microsoft degraded interoperability to competing workgroup servers because it was afraid that they may allow customers to run application on servers rather than on the Windows-dominated clients, thereby jeopardising the main source of Microsoft's power.

The Commission did not engage in a detailed analysis of the actual effects of Microsoft's conduct, but it would be difficult to imagine that it did not contribute to the dramatic changes in the market: between 1996 and 2001, Microsoft's share in the work group server market went from 20 per cent to more than 60 per cent, whereas Novell, IBM, Sun and Linux fell to about 10 per cent of the market each. The foreclosure effects may have been reinforced by network effects as noted by the Commission: the decline in market shares of the competing server OSs was likely to make it less attractive for software developers to write applications for them, which in turn would accelerate their decline.

We conclude the discussion of this case with some remarks on one of the remedies imposed: the Commission required Microsoft to disclose all the relevant information – in return for a reasonable fee – on the protocols and interfaces that were necessary for rival server OSs to fully interoperate with Windows. Microsoft complained that this compulsory licensing measure amounted to allow the rivals to appropriate its innovation and therefore had a dampening effect on innovation.

This brings us to the delicate policy issue of the trade-off arising whenever there is a compulsory licensing remedy: on the one hand, one wants to fix a competition concern; on the other hand, one should avoid unnecessary violation of protection of intellectual property rights. Two remarks in this respect. The first is that Microsoft was not required to disclose its source code, but just the information which allowed server OS rivals to have full interoperability. This arguably limited the risk that compulsory licensing led to imitation of 'real' innovative content. The second is that some of the interoperability information at issue was of arguably limited innovative content. For instance, a security protocol (Kerberos)¹²³ used in Windows originated from an open source protocol to which Microsoft included proprietary extensions which made interoperability difficult for competing servers.

¹²³ Security protocols perform crucial security functions such as authorisation and authentication.

5.6.2 Genzyme

A well-known UK abuse of dominance case is *Genzyme*. The main facts of the case, based on the assessment of the Office of Fair Trading ('OFT'), were broadly as follows. Genzyme was the only producer of Cerezyme, which at the time of the case was the only drug available for the treatment of Gaucher disease (a rare metabolic disorder): in the early 2000s only 180 UK patients were undergoing treatment. Genzyme was selling Cerezyme through the National Health Service ('NHS'). Treating a Gaucher patient with Cerezyme cost the NHS, on average, about £100,000 per year over the lifetime of the patient.

As of the date of the decision by the OFT (March 2003) there was one drug (Zavesca) which had just received authorisation to be marketed and which, the OFT considered, may have provided competition to Cerezyme going forward, although only to a limited extent.¹²⁴ TKT may have entered the market with a competing drug, but any such entry was not expected in the short-run.

Genzyme's retail price for Cerezyme charged to the NHS included the home delivery and homecare services by Genzyme Homecare. (By contrast, Zavesca was an oral drug which as such would not have required the provision of home delivery and homecare services – although as noted above, this drug was not an effective substitute to Cerezyme.)

Up until 2000 Genzyme used Healthcare at Home as its exclusive distributor and provider of homecare services for Cerezyme. Following this contract termination, Healthcare at Home, in order to continue to offer the delivery and homecare service, had to first purchase Cerezyme from Genzyme, and then agree with the NHS on a price which would have included the provision of the drug to a patient as well as delivery and homecare services.

Genzyme set the same price in both cases (£2.975 per unit or £595 per 200 unit vial): that is, it charged this amount both to the NHS for the drug including delivery and homecare services to the patient, and to Healthcare at Home for the drug only.

Healthcare at Home kept offering delivery and homecare services of Cerezyme to NHS patients past the year 2000, but complained to the OFT. The OFT found that Healthcare at Home was sustaining losses due

¹²⁴ This was due to clinical reasons, according to the OFT, including the fact that at that time Zavesca was only prescribed to patients who could not tolerate Cerezyme.

to Genzyme's pricing policy and that it was only a matter of time until Healthcare at Home would have had to leave this market.

The OFT concluded that Genzyme had engaged in an anti-competitive margin squeeze. Put in economic theory terms, this was similar to setting the downstream (or retail) price at the upstream (or wholesale) price level. This was alleged to leave no potential scope for downstream competition (that is, in delivery and homecare services). The Competition Appeal Tribunal ('CAT') confirmed the margin squeeze finding by the OFT.¹²⁵

If attempting to interpret the facts presented in this case under the economic models reviewed in this chapter, one may *prima facie* read it in the spirit of the single monopoly profit theory (or Chicago School critique) set out in Section 5.2.2: (i) selling the drug and (ii) delivering and administering it were complementary services, consumed in (near) fixed proportions. Hence, a monopolist over the drug could not increase its profits by 'leveraging' its market power over the ancillary (complementary) service.

Yet the standard Chicago argument would typically look at the market in a rather static way and would assume that the monopoly position over one market (drug supply) was exogenously set and not contestable.

In fact, the OFT noted that in addition to restricting the extent of competition in Cerezyme delivery and homecare services, Genzyme's behaviour - by preventing viable independent provision of delivery and homecare services for Cerezyme (and potentially other drugs) - also raised barriers to entry into the (upstream) market for the supply of drugs for the treatment of Gaucher disease.

[As a result of Genzyme's conduct] it is more difficult for competitors to enter the upstream market for the supply of drugs for the treatment of Gaucher disease. Since the supply of Homecare Services is effectively tied to Genzyme Homecare, a new competitor would face the additional hurdle of persuading the patient to switch not only to a new drug, but also to a new homecare services provider. (Para. 331 of the OFT's decision)¹²⁶

Expert witness testimony appeared to support this:

Professor Cox [...] expresses the view that changing homecare provider in circumstances where he was considering switching treatment could definitely affect the choice of treatment, especially in the case of vulnerable patients requiring infusion assistance, particularly since 'a very intense relationship can be built up between patients and their homecare providers'. Dr Mehta [...] also stresses that

¹²⁵ The OFT also found abusive bundling, but the CAT dismissed this.

¹²⁶ See also the summary made by the CAT at para. 480 of its judgment.

prescribing decisions have to take into account the patient's viewpoint. In Dr Mehta's view, if there is a change not just of the drug, but also of the arrangements for treatment 'from the delivery driver that he or she meets each time, to the assisting nurse with whom a relationship may have been built up and with whom the patient is content, this is not an insignificant matter'. (Para. 635 of the CAT's judgment)

There are features of the case which are consistent with the dynamic vertical foreclosure model set out in this chapter (see Section 5.2.4). However, it would have been helpful for the OFT decision or the CAT judgment to provide further information about the real chances of successful upstream entry. In its dominance analysis, the OFT emphasised the importance of upstream barriers to entry, but obviously the higher the barriers to entry, the less threatening the possibility of upstream entry and thus the less convincing the need for Genzyme to monopolise the downstream market in order to deter upstream entry.¹²⁷

5.6.3 BT/THUS/Gamma

In *BT/THUS/Gamma*, the UK communications and media regulator (Ofcom) investigated margin squeeze allegations by THUS and Gamma against BT, the leading UK fixed telecoms operator, and concluded that there was no anti-competitive conduct.

THUS and Gamma, two of BT's competitors in the supply of wholesale calls, relied on BT for some of their upstream inputs (call origination and termination services, which were subject to price regulation).

In its assessment, Ofcom distinguished between a technical margin squeeze (whereby the incumbent's downstream costs exceed the difference between upstream and downstream prices) and an actual abuse of a dominant position. The former was deemed to be a necessary, but not a sufficient condition for the latter.

¹²⁷ The CAT seemed to share the OFT's views: 'Genzyme itself saw the creation of Genzyme Homecare as a strategy which "pushes out competition, by providing a shopping basket of tailor made services". In our view, it is a reasonable inference that Genzyme considered that the creation of Genzyme Homecare would make it more difficult for competitors to Cerezyme to enter the market'. See para. 637. However, the CAT also expressed some doubts, summarising its position as follows: 'Our overall conclusion, on the balance of the evidence, is that if Genzyme were to succeed in monopolising the downstream supply of Homecare Services, that would probably have some adverse effect on the ability of a new treatment for Gaucher disease to establish itself in the United Kingdom over a reasonable timescale, but the additional foreclosure effect in the upstream market is unlikely to be as great as that suggested by the OFT in the decision'. See para. 639.

Ofcom found that an equally efficient operator to BT would have made negative margins on wholesale calls, between July 2008 and April 2009, based on the upstream and downstream prices charged by BT and on the use of long-run incremental costs identified in its analysis. However, Ofcom did not identify any resulting anti-competitive effects: it recognised, in fact, that BT's competitors continued to compete successfully for new contracts on wholesale services. Ofcom also confirmed through BT's internal documents that BT did not seem to be following any deliberate exclusionary strategy by way of a margin squeeze.

As Edwards and Walker (2013) noted, Ofcom's assessment needed to acknowledge the relevant 'arena of competition': while BT's downstream margins over the whole set of wholesale contracts were negative between July 2008 and April 2009, this conclusion was largely driven by one single large wholesale contract. Importantly, this customer was believed not to consider any supplier other than BT; if it had not found a suitable agreement with BT, it would have self-supplied (as it was doing prior to dealing with BT).

Therefore, according to Ofcom, once excluding such contract, an equally efficient operator to BT would not have made negative margins at the prices offered by BT over the time period considered in the assessment.

Ofcom thus concluded that BT's behaviour did not give rise to any anti-competitive effects.

In reaching this conclusion, it noted a number of points.¹²⁸ First, increased concentration in the market could not be conclusive evidence of an actual reduction in the intensity of competition in a market where shares can swing rapidly due to large contracts being won or lost. Second, the overall reduction in the competitors' volumes of business was part of a declining market, with large volume of business migrating to a different form of upstream access (local loop unbundling). Third, Ofcom could not find any evidence that BT's pricing strategies directly led to any market exit. Fourth, BT's competitors kept bidding for business (reseller customers) throughout the period in which BT earned a negative margin on its wholesale calls. Fifth, the costs of BT's competitors did not change significantly in spite of a fall in volume and Ofcom could not identify any change in the shape or size of their networks. Sixth, Ofcom noted that BT's competitors appeared to have spare 'TDM capacity',¹²⁹ so that

¹²⁸ See paras 7.227–7.232 of the decision.

¹²⁹ Time-division multiplexing ('TDM') is a way of transmitting and receiving signals through synchronised switches at each end of the transmission line.

this (rather than BT's conduct) limited their appetite for further rollouts (another reason, Ofcom noted, was their focus on investing instead on new generation internet protocol networks).

We next move to a possible interpretation of the facts presented in this case using the economic models reviewed in this chapter.

We noted above the existence of a large contract that affected Ofcom's assessment of the alleged margin squeeze by BT. Ofcom also noted that there was a second large contract, which may have been contested by one or two competitors. In that context, Ofcom stated:

[T]here are circumstances in which individual contracts may be of particular importance to competition, and therefore, the margins earned on such contracts are of interest. For example, in markets where there are strong economies of scale, individual contracts could be considered to be 'competition enablers'. In such circumstances, we would be concerned about pricing below cost on such contracts even if the product as a whole was profitable. This is because winning the contracts that drive the largest volumes will be important for firms to enable them to operate at, or beyond, the minimum efficient scale. Also, even if an incumbent or dominant firm is operating significantly beyond the minimum efficient scale it may adopt a strategy of aggressively pursuing large volume contracts as a mechanism to deny entrants the opportunity to reach the minimum scale and therefore effectively compete in the market. (Para. 6.42, footnotes omitted)

That is, Ofcom was open in principle to a theory of harm based on the general mechanism discussed in this book. Had this theory of harm been backed by evidence (including of anti-competitive effects), the margin squeeze that Ofcom may have found could have been interpreted in the spirit of the model presented in Section 5.3.2.2: put simply, in the presence of price regulation which limits the rents that the incumbent can extract from downstream rival, the incumbent could have profitably and successfully engaged in margin squeeze. The rationale for the margin squeeze, under that scenario, would be to exclude a competitor by denying it scale (or a range of customers, for example) in the short-run, forcing it to exit the market, or relegating it to a fringe, in the medium- to long-run.

On the facts of the case, however, Ofcom noted that BT's potential competitors for the second contract already operated at high levels of interconnection at digital local exchanges (that is, could already reap significant scale economies), so that any incremental volumes would have been unlikely to reduce average costs to any material extent. In other words, a theory of harm based on the strategic denial of scale or scope economies was unwarranted in this case, according to Ofcom. We do not have enough independent information on the case to judge whether Ofcom was right in

its conclusion, but we note this as a possible example of a case where an exclusionary mechanism based on the denial of scale or scope economies was considered but discarded on the basis that the facts of the case did not support the theory.

5.6.4 Deutsche Telekom

In *Deutsche Telekom*, the Commission found that the eponymous German incumbent operator had engaged in margin squeeze in the market for local access to telephone networks between 1998 and 2003. In particular, the Commission found that Deutsche Telekom ('DT') left a negative margin over a period of three years between the charge for unbundled access to its local loop (that is, the 'wholesale' or 'upstream' price) and the retail access charge to end-users. The Commission imposed a fine on DT of €12.6m. Both the General Court and the Court of Justice upheld.

DT was subject to regulation by the national sectoral regulator (then called RegTP). At the wholesale level, since 1997, DT had to offer access to the unbundled local loop ('ULL', through which downstream competitors could offer any retail service), on a cost-oriented basis; this price level was regularly reviewed by the RegTP. In addition, at the retail level, DT was subject to a retail price cap on analogue (traditional) telephony and ISDN (narrowband connections), but not on broadband connections (which could rely on either analogue or ISDN connections). The retail price cap operated as a 'price cap index', that is a maximum charge for a basket of services that included both retail access and call charges (rather than a price cap on each retail service). The level of this price cap index was being reduced over time.¹³⁰

The Commission found that DT had a monopoly in the market for local network access at the wholesale level and had a market share of 94 per cent in the retail broadband access market (its share in the retail narrowband access market was higher).

The Commission carried out the 'margin squeeze test'. It stressed that it was important to consider comparable products at the wholesale and the retail level in a margin squeeze investigation (that is, ULL access charges to competitors *versus* retail access charges to end-users) and thus excluded call revenues for its calculations.

¹³⁰ The RegTP imposed a reduction of 4.3 per cent in the period from 1 January 1998 to 31 December 1999 and a further reduction of 5.6 per cent in the period from 1 January 2000 to 31 December 2001.

The Commission noted that DT reduced its retail charges on calls more than the reduction in the price cap index would have envisaged. Possibly as a result of this, the market shares of DT's competitors of analogue connections fell from 21 per cent in 1999 to 10 per cent in 2002, according to the Commission.

The Commission took care to explain how DT could have abided by the regulator's prescriptions *and* have acted in accordance to competition law provisions. This would have involved tariff-rebalancing. First, DT should have raised retail broadband charges (as these were not regulated, that is, the price cap index excluded them, as noted above). Second, having reduced its call charges faster than the glide path indicated by the RegTP, DT should have increased its retail access charges. In doing so carefully, DT could have ensured that (i) its weighted average retail charges would have remained below the price cap index; and (ii) an as-efficient competitor purchasing ULL from DT could make a non-negative margin on the access service alone (that is, excluding calls).

On the other hand, the Commission conceded that DT was barred by the RegTP from raising the retail narrowband access price after January 2002. The only way DT could avoid a margin squeeze after this date would have been to raise retail broadband prices. DT defended itself by stating that this was a nascent market and introductory offers were necessary to stimulate demand but this left the Commission unconvinced.

We find it quite difficult to make economic sense of this antitrust decision, possibly because the Commission did not set out a coherent theory of harm (the Commission's assessment was validated by both EU Courts in a fairly formalistic way). We limit ourselves to three broad remarks.

First, as we argued in Section 5.4.2 we believe that in industries where there is a sectoral body in charge of economic regulation (including issues around access conditions) any antitrust enforcement should only take place under exceptional circumstances. The goals of economic regulation may include fostering markets where end-consumers can enjoy high quality services, competitive prices and efficient levels of investments in infrastructure. We do not think that the objectives of competition policy are dramatically different. Yet, while competition law may – legally – have a superior stance in the EU than *ex ante* regulation, applying competition law in markets that are subject to sectoral economic regulation may end up seriously undermining legal certainty for businesses.

Second, as we discussed in Section 5.4.2, competition authorities may want to be particularly careful before intervening in nascent markets. One may argue that broadband services were at their infancy in the late 1990s and early 2000s and that a telecoms operator needed to provide appealing introductory offers to generate a viable customer base. The fact that the Commission suggested that DT should have raised retail broadband prices at that point in time strikes us as possibly being incompatible with the promotion of consumer welfare.

Third, if we were to engage more closely with the Commission's actual analysis of this case, it is arguable whether the Commission was right to exclude revenues from retail calls from its assessment of the alleged margin squeeze.¹³¹ Put otherwise, the Commission may have adopted the wrong aggregation level for its test. The ULL service gave competitors the ability to sell both access and call services (which were retail products also offered by the vertically integrated operator). And indeed, as the General Court also accepted, 'from the point of view of the end-user, [retail] access services and call services constitute a whole' (para. 199 of the judgment). It may therefore have been more appropriate for the Commission to include retail call revenues in its assessment, even if this would have made the assessment more involved (given a local loop connection was not associated with a pre-determined level of calls or data).

5.6.5 Telefónica

In *Telefónica*, the Commission found that the eponymous Spanish telecoms incumbent abused its dominant position by way of a margin squeeze in the Spanish broadband market, from September 2001 to December 2006.

According to the Commission's decision, the facts of the case were broadly as follows. Telefónica was the unique operator having a local access network, that is, a network that reaches final users. Alternative operators wishing to provide services throughout Spain had no other option than buying wholesale services from Telefónica. Access to wholesale services could occur in three different ways, which differed in the intensity of the investment required by an alternative operator to supply the retail market: (i) *Unbundled local loop* ('ULL'), which allowed alternative operators to use only the very final portion of the incumbent's network,

¹³¹ See O'Donoghue and Padilla (2013) for a further critique, stressing that most retail offers by incumbents and their rivals alike include bundles of access and call or data services.

but required them to make a significant investment in their own network. This solution allowed alternative operators to gain discretion in how to provide retail services and the ability to differentiate their services from Telefónica's. (ii) *Regional wholesale access*, which required alternative operators to invest less as compared to the case of ULL, although the investment was still substantial. Alternative operators would have less discretion in the provision of retail services. (iii) *National wholesale access*, whereby alternative operators would require relatively limited network investment, but would not have much discretion in the provision of retail services.

In this context, the Commission identified as the relevant retail market the broadband services to the 'mass market' of residential and non-residential users; and it identified as relevant (separate) wholesale markets the broadband access at regional and national levels, respectively.¹³² The Commission considered that access through ULL was not in the same market as national and regional wholesale access, because in its view switching from national and regional access to ULL would be very costly and time-consuming. The margin squeeze investigation therefore disregarded ULL and focused on a comparison between wholesale and retail broadband access prices.¹³³

Telefónica was found dominant in all of these markets: it was found to be a monopolist in the wholesale regional access market and to have over 84 per cent of the wholesale national access market. At the retail level, Telefónica's share of end-users ranged between 52 and 58 per cent during the period of the abuse found (its market share by revenue consistently exceeded 60 per cent according to the Commission).

Looking at the regulatory environment, Telefónica's retail prices were regulated up until November 2003.¹³⁴ Following that date, retail prices were liberalised. As for wholesale products, national wholesale access was not price-regulated during the period of the infringement. Regional wholesale access was regulated by price caps set on a 'retail-minus' basis (that is, Telefónica could not charge for wholesale more than (100-x) per cent of what it charged at the retail level for the same product,

¹³² According to the Commission, whether wholesale regional access was in the same market as wholesale national access would not have changed the conclusion of its assessment.

¹³³ Telefónica, instead, argued that ULL should be in the same market as regional and national wholesale access, since they allowed alternative operators to provide the same retail services.

¹³⁴ Telefónica had to propose a retail price to the regulator for approval.

where x was the proxy set through regulation for the incremental cost of providing downstream services and any network elements sought by the access-seeker).¹³⁵

Telefónica's retail prices remained fixed (in nominal terms) throughout the whole period from September 2001 to December 2006; however broadband speeds were upgraded for each retail service class (for example, basic package, premium package).¹³⁶

To establish whether there was a margin squeeze, the Commission applied the usual test: it checked whether the margin between Telefónica's retail price and wholesale price would allow an equally efficient competitor to cover the downstream long-run average incremental costs ('LRAIC'), that is, to cover the average additional costs that an operator has to incur to operate downstream. The measure of LRAIC used in the investigation included costs for additional network elements needed to provide retail services, recurrent costs of internet service providers, customer acquisition costs (advertising, incentives and commission to the sales network) and a share of common costs (for example, costs associated with the commercial or company structure).¹³⁷

In terms of the aggregation level chosen, the Commission carried out the margin squeeze test using Telefónica's actual mix of retail products, and

¹³⁵ One may wonder how margin squeeze could occur given that the wholesale price was determined using a retail-minus system. The Commission argued that regulated prices had been determined in 2001 on the ground of *ex ante* estimates of costs that were eventually lower than the costs actually incurred. Moreover, it argued that the price indicated by the Spanish regulator was a maximum price that Telefónica was free to decrease. Hence, Telefónica had the duty to check whether the combination of the retail prices and regulated price was replicable by an equally efficient competitor and to take into account that the forecasts used *ex ante* to determine the regulated wholesale price were not confirmed by the actual evolution of the market. One may therefore read the proceedings against Telefónica as the Commission 'making up' for an incorrect regulation. The regulator started imposing cost-oriented offers on both wholesale products from 2007.

¹³⁶ The Commission (borrowing the remark from the Spanish regulator) also noted how a 'key competition tool' in the retail market were promotional offers, such as gifts or discounts (or waivers) of connection fees, subscriptions fees and equipment fees. Likewise, bundled offers started to emerge in the market: 'double play' (fixed line and internet) and 'triple play' (which also included TV over broadband). The former was offered by both the incumbent and the competitors, the latter mostly by Telefónica, since competitors lacked the necessary infrastructure to provide it nationwide.

¹³⁷ The details of the exact determination of the LRAIC were the object of involved debates. For instance, Telefónica argued that the costs related to the commercial structure should be excluded from the LRAIC.

not at an individual product basis. This seems to us as more reasonable than the approach in *Deutsche Telekom*, the case just discussed.¹³⁸

To assess the viability of an as-efficient competitor, the Commission engaged in two profitability tests: first, it sought to establish whether the difference between the incumbent's retail and wholesale prices allowed to cover (downstream) costs, on an annual basis. Second, it applied discounted cash flow analysis, considering total discounted revenues and costs over the period of the abuse. According to the Commission, the latter methodology allowed to account for the fact that the provision of a new service, such as broadband services, might entail initial losses.¹³⁹ Using either test, the Commission concluded that a margin squeeze had occurred, and imposed a fine of €151.9m on Telefónica. On appeal by Telefónica, the General Court and then the Court of Justice upheld the Commission's decision.¹⁴⁰

We next discuss the Commission's economic assessment of this case in more detail.

In the implementation of the test, the Commission separately compared the margin between Telefónica's retail and wholesale prices to the downstream costs of an equally efficient competitor relying *entirely* on national wholesale access services and to those of an equally efficient competitor relying *entirely* on regional wholesale access services. Telefónica objected that this would not have been the profit-maximising strategy for an entrant: entrants would typically cherry-pick customer categories and regions and would choose an optimal mix of regional access, national access and of

¹³⁸ The Commission noted: 'In the case at hand, the margin squeeze test has been conducted on the basis of an aggregated approach, i.e. on the basis of the mix of services marketed by Telefónica on the relevant retail market. This approach (referred as to the 'aggregated approach') is based on the principle that competitors must at least be able to profitably replicate Telefónica's product pattern. This is the approach most favourable to Telefónica, since it gives it maximal flexibility to spread the costs which are common to its retail products (provided that the margin squeeze test yields a positive result with the aggregated approach). The aggregated approach is consistent with a new entrant's internal decision making process in that it assesses the profitability of its investment in a network by considering the complete range of products that it is able to offer in the relevant downstream market' (para. 388 of the decision).

¹³⁹ The Commission expressed some reservations about the latter methodology. In particular, it noted that a positive net present value is consistent with the incumbent making short-term losses and large long-term gains, achieved through exclusionary behaviour and an increase in market power.

¹⁴⁰ The Kingdom of Spain also appealed the Commission's decision to the General Court, on the grounds that the Commission violated its duty to cooperate with the Spanish regulator and that it breached the principle of legal certainty, given that Telefónica was subject to sectoral regulation. However, this appeal too was dismissed.

ULL. The Commission justified its approach with reference to how other markets evolved in other countries. In particular, the Commission pointed out that the most profitable entry strategy for an alternative operator was to invest in its own network, combined with ULL use:

[...] due to the risks involved in such a high sunk investment, alternative operators are likely to follow a step-by-step approach to continuously expanding their customer base and infrastructure investments. When climbing up the ‘investment ladder’ [...], alternative operators seek to obtain a minimum critical mass, in order to be able to make further investments. [...]

The first step of the ‘investment ladder’ is occupied by an operator whose strategy consists in targeting a mass market (thus involving considerable marketing and advertising expenditure), but who is merely acting as a reseller of the ADSL access product of the vertically integrated provider (the incumbent). As its customer base increases, then the alternative operator makes further investment. In a further step, it may even seek to connect its customers directly (local loop unbundling). Thus the progressive investments take the alternative operator progressively closer to the customer, reduce the reliance on the wholesale product of the incumbent, and increasingly enable it to add more value to the product offered to the end-user and to differentiate its service from that of the incumbent. (Paras 392 and 178 of the decision)

This argument may explain why the Commission decided to perform a margin squeeze test separately for regional and national wholesale access:

It is therefore necessary that there should not be any margin squeeze in relation to any step of the ladder, i.e. in relation to any wholesale product. If there was such a margin squeeze, new entrants that are climbing the ladder of investment would be foreclosed. (Para. 392 of the decision)

If the Commission’s reasoning in relation to the investment ladder was correct, it would provide a possible rationale for Telefónica’s incentive to engage in margin squeeze, which may be consistent with the dynamic theory we have illustrated in Section 5.2.4 and in technical Section 5.3.6. Using that lens, only if an alternative operator (relying on national or regional access) obtained a critical size in the retail market, would it subsequently be able to make the investment necessary to reach customers directly through ULL and to rely less on the services provided by the incumbent.¹⁴¹ By engaging in margin squeeze, the incumbent would prevent alternative operators from achieving the critical size that would justify investment in their own infrastructure, thereby discouraging them from investing further upstream. Margin squeeze may therefore potentially

¹⁴¹ For the ladder of investment theory, see Cave (2006) and Bacache et al. (2014).

be interpreted as a defensive strategy adopted by the incumbent to protect its position in the upstream market.

As a part of the assessment of the possible incentives to exclude, one should also assess whether, in the specific market involved, a margin squeeze strategy is likely to lead to the exclusion of the downstream rivals. On the one hand, Telefónica was found to have a dominant position in the retail market which, as we discussed in Section 5.4.2, usually makes it more likely that margin squeeze would lead to exclusion. On the other hand, the broadband market was a growing market at the time of the decision. As we discussed in Section 5.4.2, exclusion of downstream rivals is less likely to be successful when future demand is expected to grow. Moreover, in a market at an infant stage, low retail prices are more likely to be justified by the intent to make consumers familiar with the new product and services (see Section 5.4.2.2, on ‘penetration pricing’). Furthermore, some alternative operators were affiliates of international telecoms groups (such as France Télécom and Deutsche Telekom). One may argue that competing operators with potentially significant financial backing may be in a stronger position to react to Telefónica’s offers (and potentially sustain losses for a period) than otherwise, in order to achieve a critical mass of users. This may undermine any incentive by Telefónica to engage in margin squeeze to exclude. In the decision, the Commission noted that ‘only companies with a sufficiently strong financial backing have been able to survive and grow (slightly, and at a loss) in the mass market’ (para. 587). These (two) operators, however, did not expand significantly and probably did not achieve a critical mass of users at least in the short-term, while none of the other operators achieved a market share of 1 per cent or more during the period of the Commission’s assessment.

5.6.6 Long-term Capacity Bookings (RWE, GDF, E.ON)

The European Commission’s inquiry of the European energy sector,¹⁴² which lasted between 2005 and 2007, resulted in a number of investigations. Here we focus on one of the Commission’s key concerns arising from this sector inquiry, that of long-term capacity bookings. We refer to three prominent cases, *RWE*, *GDF* and *E.ON*, which all ended with undertakings

¹⁴² See European Commission, ‘Competition: Commission opens sector inquiry into gas and electricity’, Press Release IP/05/716, 13 June 2005; and ‘Competition: Commission energy sector inquiry confirms serious competition problems’, Press Release IP/07/26, 10 January 2007.

offered by the incumbents to address the Commission's competition concerns.^{143,144}

These cases appeared to share a similar background, according to the facts set out by the Commission: within their geographic area of operation, each incumbent had a subsidiary which was an upstream monopolist (owning the gas transmission network) and a downstream subsidiary (the gas sales' business), which faced competition.

Long-term capacity bookings consisted of the upstream and the downstream subsidiaries agreeing on long-term contracts granting the downstream affiliate most of the network capacity of the upstream affiliate. Downstream competitors approaching the upstream division of the incumbent with requests for capacity were told there was none or little available.

This, the Commission's allegations went, led to very limited downstream competition. The Commission was also concerned that the long-term nature of such contracts may also have damaging effects on potential competition and in particular on the incentives for downstream competitors to enter in the medium-term, expecting a shortage of upstream capacity.

In all these cases, the Commission considered the transmission network an 'essential facility' (see our discussion in Section 5.2.1).

RWE In the case of RWE, the Commission raised concerns about the foreclosure of RWE's downstream competitors: by booking 'almost the entire capacities on its transmission network on a long term basis' for its downstream subsidiary, RWE limited access to its gas network to downstream competitors (this would amount to a refusal to supply, according to the Commission).¹⁴⁵ The Commission was also worried about a margin squeeze to the detriment of RWE's downstream competitors, evidenced, according to the Commission, by 'RWE's elevated prices for access to its transmission network' and by RWE's consistently negative margins downstream. Moreover, according to the Commission, RWE was offering substantial rebates (quantity discounts) at the wholesale level. However, in practice, it was almost exclusively RWE's downstream subsidiary that benefitted from those rebates, mainly because its rivals did

¹⁴³ We discuss the long-term contracts in *Distrigaz* among the cases reviewed in Chapter 3, in the context of exclusive dealing.

¹⁴⁴ In what follows, we also draw from Federico and Vives (2008), as well as from material prepared by Giulio Federico in the context of graduate teaching at the Barcelona Graduate School of Economics, for which we thank him.

¹⁴⁵ Formally, the Commission did not challenge the long-term nature of these contracts *per se*, but this is secondary for our discussion here.

not obtain sufficient capacity to attain the relevant quantity thresholds, according to the Commission.

Further, the Commission found that RWE discriminated against downstream competitors by charging very high penalty (balancing) fees (which its downstream arm was not subject to).¹⁴⁶

Eventually RWE offered to divest (with a minor exception in the area around Bergheim) its entire German gas transmission network. The Commission accepted this undertaking and closed the investigation in 2009.

GDF The facts presented in *GDF* were similar.¹⁴⁷ The upstream affiliate (GRTgaz) had originally booked (prior to the Commission's on-site inspections in 2006) the vast majority of its transmission capacity (at the main entry points) for GDF's downstream subsidiary until 2019. GRTgaz owned and operated all the main gas import entry points. Moreover, another GDF subsidiary, Elengy, owned and operated the two French liquefied natural gas ('LNG') terminals.

The key concern by the Commission was that GDF's long-term contracts would foreclose gas import capacity in each of the balancing zones of the GRTgaz network, with a resulting restriction of competition in the markets for the supply of gas in those zones. Second, the Commission was concerned about its preliminary finding that GDF had allocated a very significant proportion of total capacity at the Fos Cavaou LNG terminal to its own gas trading division and to Total, without any open, transparent and non-discriminatory procedure. Third, the Commission was concerned about GDF's possible underinvestment in additional import capacity at its Montoir de Bretagne LNG terminal (we discuss 'strategic underinvestment' further when reviewing *ENI* in Section 5.6.7).

The undertakings offered by GDF (and accepted by the Commission) were less drastic than in *RWE*, in that behavioural remedies were deemed to be sufficient to address the Commission's concerns, without resorting to structural remedies. In particular, GDF agreed to cap its booking share at 50 per cent of its H-gas (high-calorific) network capacity by 2014, with immediate release of significant capacities (worth around 10–15 per cent of total capacity).

¹⁴⁶ As the Commission put it, '[b]alancing services are intended to bridge the differences between forecasted and actual transport volumes, i.e. the balancing service provider buys gas from shippers if these have unexpected excess capacities and sells gas to shippers if they need more gas than expected' (see footnote 30 of the decision).

¹⁴⁷ See Cardoso et al. (2010) for a discussion of this case, presenting a few more facts than the Commission's decision.

E.ON *E.ON* was a similar case, as presented by the Commission, although it concerned the transmission capacity of national gas as opposed to import capacity. *E.ON* was found to be the largest supplier of natural gas and the largest gas transmission network operator in Germany. It operated in the German gas market via its fully owned subsidiary *E.ON Ruhrgas*, which itself owned the upstream transmission network operator *E.ON Gastransport*.

The product markets identified were H-gas and L-gas (high- and low-calorific gas, respectively). *E.ON* was found to have a dominant position in the supply of both types of gas downstream (with market shares of around 60 and 80 per cent, respectively, and with a stronger position in the case of industrial customers). *E.ON* was also deemed to be an upstream monopoly in the case of L-gas transmission and to control about 80 per cent of the market in the case of H-gas. *E.ON* had long-term booking contracts (until at least 2019) with its downstream affiliate covering around 80 per cent of its H-gas freely allocable capacity, and in excess of 90 per cent in the case of L-gas.

The undertakings offered by *E.ON* (accepted by the Commission in 2010) included an immediate release of capacity at key entry points (amounting to about 15 per cent of capacity), with a commitment to cap, by October 2015, *E.ON*'s H-gas booking share at 50 per cent (and 64 per cent in the case of L-gas).¹⁴⁸

Some remarks We conclude this section with a brief discussion of policy considerations in relation to the long-term capacity bookings just discussed.

First, in none of these cases did the European Commission articulate a theory of harm. Possibly this is due to the fact that these are commitment decisions, and as such the Commission can limit itself to set out its preliminary concerns. But it would nonetheless be interesting to understand what the anti-competitive rationale behind long-term capacity bookings is (on the basis of the information in the public domain, we do not see any of the theories of harm analysed in this chapter to be naturally applicable to these cases), and to assess whether one can discard possible efficiency defences.

Second, and related to the previous point, it is difficult to see how the objective to exclude downstream rivals may co-exist with excessive

¹⁴⁸ As noted in a previous footnote, the Commission released *E.ON* from its commitments almost five years ahead of schedule, as it believed that competition in the market had increased significantly. See European Commission, Press Release IP/16/2646 of 26 July 2016.

upstream prices. In *RWE*, for instance, the Commission appears concerned with a possible margin squeeze implemented through high access prices. But in such a situation, the vertically integrated firm would not have an incentive to exclude downstream rivals, as it would be free to extract rents from them. Furthermore, if price regulation is deemed to be ineffective and is thought to allow for too high upstream prices, then structural separation would not solve the underlying issue, as the upstream subsidiary would continue to charge high prices.

Finally, from a policy perspective, it is arguable whether antitrust intervention is truly warranted in a highly regulated industry such as the energy one. To the extent that antitrust intervention aims at remedying any form of *ex ante* regulation which is deemed imperfect, we wonder whether it would be more effective to fix instead the regulatory regimes. In a sense, this may be what the European Commission claims to have embarked on, through its so-called Third Energy Package, aimed at making the EU energy market ‘fully effective’, at creating a single EU gas and electricity market and at increasing standards of service and security of supply.¹⁴⁹

5.6.7 ENI

The facts in *ENI* were somewhat different from the long-term capacity bookings just considered. In 2006, the European Commission started raising concerns in relation to some behaviour by ENI, the Italian incumbent oil and gas operator. The Commission alleged that ENI: (i) refused to grant competitors access to capacity available on the transport network (capacity hoarding);¹⁵⁰ (ii) granted access in an impractical manner (capacity degradation); and (iii) strategically limited investment in its international transmission pipelines into Italy (strategic underinvestment).¹⁵¹

¹⁴⁹ See Directives 2009/72/EC and 2009/73/EC and Regulations (EC) 713/2009, (EC) 714/2009 and (EC) 715/2009, as well as European Commission, MEMO/11/125 of 2 March 2011. All EU Member States had to transpose these two Directives into national legislation by 3 March 2011. An important element of the Third Energy Package was the unbundling of energy production and supply interests from the networks. Unbundling does not necessarily require vertical separation (for example, a supply company can retain ownership of the network but this must be operated and maintained by an independent company, which would also be responsible for the investments in the network). Further, third party access to the networks would have to occur through regulated tariffs and on a non-discriminatory basis.

¹⁵⁰ Maier-Rigaud et al. (2011) stated that ENI’s competitors would obtain on average less than 3–10 per cent of the available capacity on the pipeline.

¹⁵¹ These pipelines were the Trans Europa Naturgas Pipeline and Transitgas pipeline, conveying North European gas through Germany and Switzerland, and the TAG

These actions, the Commission argued, coupled with ENI's position in the upstream and downstream markets,¹⁵² may have amounted to an abuse of a dominant position in the gas transportation market into Italy, to the detriment of downstream competitors and, ultimately, consumers, in the Italian gas supply market. In 2010, to address the Commission's concerns, ENI offered to divest its holdings in the companies running those international pipelines. The Commission believed that this undertaking was both necessary and sufficient to solve its competition concerns.

Put into economic terms, the Commission's original findings may be interpreted as a vertically integrated incumbent restricting upstream capacity so as to reduce the ability of downstream competitors to compete effectively in the marketplace. Ultimately, the Commission reasoned, this would lead to harm to end-consumers.

In this sense, although the main allegations in *ENI* appeared different from the other energy cases reviewed in Section 5.6.6,¹⁵³ the Commission's ultimate concern – in economic terms – was the inability for downstream competitors to get access to sufficient upstream capacity in order to effectively compete downstream.¹⁵⁴

Maier-Rigaud et al. (2011)¹⁵⁵ elaborated on the Commission's thinking set out in the public version of its decision and stressed the presence of market share caps on gas inflows imposed by Italian law on ENI: during the period from 2002 to 2010 no operator was allowed, in a given year, to import or produce more than 75 per cent of domestic gas consumption (this cap was progressively reduced by 2 percentage points each year down to 61 per cent). The authors conjectured that this constraint limited ENI's

pipeline, from Russia. According to the Commission, the gas imported via these pipelines amounted to over one half of total gas imports to Italy. Further, also according to the Commission, the limitation of upstream investments was not driven by the lack of profitability from increased transportation activity for ENI as a transmission system operator but was rather 'meant to protect ENI's own downstream profits at the detriment of profits on the transportation level in order to maximize overall profits' (see para. 59 of the decision).

¹⁵² The Commission referred to data indicating that ENI's share of domestic production was around 85 per cent in 2007, while its share of imports was around 65 per cent, or up to 70 per cent, if including ENI's sales of gas to national suppliers at the Italian borders.

¹⁵³ The Commission, in its decision, did however also refer to 'long term bookings by ENI of a significant part of the existing [import] capacity' (see para. 33 of the decision).

¹⁵⁴ From a more formal economic modelling perspective, in *GDF, RWE and E.ON* upstream capacity was – by and large – taken as fixed by the Commission, while in *ENI* the Commission argued that upstream capacity was actively being limited.

¹⁵⁵ As noted in this article, all three authors used to work at the European Commission's Directorate General for Competition, Unit B-1, Energy and Environment Antitrust.

ability to expand its downstream market share in reaction to any price competition, so that ENI's strategy

consisted of maintaining and securing its supply margins by preventing the development of effective competition in the downstream markets...To protect its profits downstream, ENI retained control over the transport routes, by embarking upon a strategy of deliberately keeping capacity tight in order to limit third parties' access to import infrastructures and therefore foreclose downstream gas supply markets.

In this sense, any exclusionary behaviour by ENI could not be 'exclusionary' in the strict meaning of the term; rather, at least according to Maier-Rigaud et al. (2011), ENI's objective was to keep the extent of competition approximately at a set level (in market share terms, rather than in absolute sales terms), while maximising the margins on its own sales.

As for the limb of the Commission's original allegations against ENI on 'strategic underinvestment', we are somewhat sceptical about adapting a refusal to supply framework to a 'refusal to invest' framework, for at least two reasons.

First, it would be difficult (especially for a competition authority without sectoral expertise) to determine what the 'right' (or non-abusive) level of investment in upstream infrastructure is.

Second, the assessment would be complicated by the existence of an opportunity cost of that investment (that is, the net return from the next-best alternative that the incumbent could have invested in), which would have to be examined in detail.

Finally, we note that, as part of its discussion of 'capacity degradation', the Commission alluded to alleged behaviour by ENI aimed at staggering the sales of capacities from complementary pipelines in a way that would reduce the value to a bidder:¹⁵⁶

It also appears that allocation procedures were designed in such a way that they would result in separate and uncoordinated capacity sales on complementary pipelines (such as TENP and Transitgas). This lack of coordination may have discouraged or prevented shippers from obtaining capacity they initially had, or had intended to, bid for. (Para. 53)

Potentially, one way of interpreting this suggestion by the Commission is that the staggering of the capacity sales reduced the extent to which bidders could realise scale and scope economies. If that were the case, this may be seen as a possible first step towards establishing a broader evidence

¹⁵⁶ See para. 53 of the decision.

base underpinning a potential theory of harm based on denying scale or scope economies to competitors, with a view to restricting the extent of competition in the future. However, given the very short discussion of the relevant fact in the Commission's decision, it is not clear at all whether this plank of a possible theory of harm was sufficiently well borne out by the facts in *ENI* and even less so if *ENI* acted with this specific strategy in mind.

To conclude, it is arguable whether antitrust intervention is appropriate in a highly regulated industry. Further, especially in such a context, adopting theories of harm such as the ones put forward in *ENI* may be quite controversial, and at the very least they should be better substantiated.

5.6.8 Google (comparison shopping)

On 27 June 2017, the European Commission fined Google € 2.42 billion for abusing its market dominance as a search engine by promoting its own comparison shopping service in its search results, and demoting those of competitors.¹⁵⁷

On the basis of the Commission's assessment, there are two relevant markets at issue in this case. One is the search engine market, in which Google – which has more than 90 per cent of market queries in Europe – is deemed to be dominant. The other is the European market for comparison shopping, namely the service which allows consumers to compare products and prices and make online purchases.

Google entered the latter market in 2004, with a product that was initially called 'Froogle', and then re-named 'Google Product Search' in 2008. Since 2013 it has been called 'Google Shopping' and has introduced some new features: in particular, consumers clicking on cells containing pictures and prices of products would go directly to a merchant's website (making it more similar to an advertisement model which is directly monetised by Google), whereas in the previous versions of the service, clicking would bring the consumer to Google's comparison service.

¹⁵⁷ When the *Google (comparison shopping)* decision was taken, this book was already in production. Here we briefly discuss the case on the basis of the few public statements of the Commission (see in particular Press Release IP/17/1784 of 27 June 2017) – as the full decision is not available at the time of writing – as well as remarks made throughout the case by Google and complainants, with the aim of explaining the main features of the case. Massimo Motta was Chief Competition Economist during an earlier period of the *Google (comparison shopping)* investigation, and hence did not participate in the drafting of this subsection.

When Google entered comparison shopping markets with Froogle, there were already a number of established players, and reportedly Google was unsatisfied with Froogle's market performance relative to its comparison shopping rivals.

According to the Commission, Google decided in 2008 to rely on Google's dominance in general internet search to improve its performance in the comparison shopping markets. In particular, Google has since then resorted to a strategy which involves two elements, according to the Commission. First, it has systematically given prominent placement to its own comparison shopping service: when a consumer enters a query into the Google search engine, Google's own comparison shopping service results are displayed at or near the top of the search results. Second, Google has demoted rival comparison shopping services in its search results. As a result of changes in Google's generic search algorithms (incorporated in the so-called 'Panda' version), rival services have started to be ranked very low, and hence virtually disappeared from a Google search page.¹⁵⁸ Moreover, Google's own comparison shopping service is not subject to Google's generic search algorithms, including such demotions.

As a result, Google's comparison shopping service is much more visible to consumers in Google's search results, whilst rival comparison shopping services are much less visible (users tend to focus on the very top results – and this effect is more pronounced in mobile devices because of the smaller screen size – and very rarely browse beyond the first page of generic search results).

This preferential treatment to own comparison shopping services and the simultaneous demotion of rival services, has according to the Commission distorted traffic (namely, consumers' visits) away from competitors and towards Google's own services. Traffic is crucial in this line of business for two reasons: first, because it leads to more clicks and hence to more revenue; second, because online retailers would be willing to list their products in a comparison shopping website only to the extent that it reaches enough consumers (that is, it generates enough traffic).

Which theory of harm? Given that the Decision has not been published yet at the moment of writing, we can only make conjectures about the precise theory of harm followed by the Commission. However, from the

¹⁵⁸ The Commission states that there is evidence that even the most highly ranked rival service appears on average only on page four of Google's search results.

information available it seems to us that this case may potentially be assessed as a tying or as a vertical foreclosure case.

Under the tying interpretation, Google would be tying its Google shopping service pages to the pages featuring the results of Google's organic search whenever the consumer queries are related to shopping. In other words, Google would leverage its dominance in the organic search market into the comparison shopping market through tying.

Under the vertical foreclosure interpretation, Google would be dominant in the upstream market of organic search services and would effectively deny access to such an input to its comparison shopping rivals, while giving prominence to its own services.

Both in the case of tying and of vertical foreclosure, we know that the Chicago School has pointed out that exclusionary strategies may well be unprofitable, so we should ask what the anti-competitive rationale behind Google's strategy would be.

One particular feature of these markets which may play an important role is given by the fact that these are two-sided markets in which consumers (as well as the sites which are ranked by the Google search engine) pay a zero monetary price (as users, we are not charged for either browsing general search results or for visiting comparison shopping services) while it is only retailers/merchants which will pay Google (or its rivals) for the consumers' clicks. Therefore, the typical Chicago School critique that exclusion would not be profitable because a dominant firm would make more money by appropriately pricing its products or services does not apply here: given its business model, Google does not charge consumers, nor comparison services for inclusion in its search pages. The monetisation of its search services would come only when consumers go to its 'Froogle' or 'Google Product Search' webpage and click on a merchant post; or when they click on the image of a merchant in its 'Google shopping' service.

Therefore, in a world where consumers would go to Google's search page to look for price and product comparison, and from there they would go to a rival comparison shopping website, Google would not obtain any revenue from the comparison shopping market (since it does not charge rival comparison shopping services for inclusion in the search pages). It may just obtain revenue if consumers clicked on some ads placed in the general search page, but the latter may not be a frequent event if consumers got to the search page just to compare online offers. Furthermore, to the extent that over time consumers may start to like some particular comparison

shopping site, they may even bypass the Google search page completely and go directly to their favourite website.

Therefore, a possible theory of harm is that – by demoting rival comparison shopping sites and placing its own in a prominent way – Google managed to obtain revenues that otherwise would have not received, as well as – in a more dynamic perspective – to avoid the risk that a number of consumers would eventually not resort to the Google search page as a first step in their process of comparing online offers.¹⁵⁹

A crucial issue is whether Google's conduct has had anti-competitive effects, namely it has harmed consumers or just some competitors. Let us start with effects on competitors. Since the decision is not available at the time of writing, we can only rely on the data posted in blogs by Google and complainants.

Effect on competitors. Google has always denied that its conduct may be anti-competitive and, rather, has claimed among other things the following:¹⁶⁰

(1) 'There is more choice than ever before'. Google argues that there are numerous other search engines, as well as 'a ton of specialized services like Amazon, Idealo, Le Guide, Expedia or eBay', and that people use social sites and mobile phone apps to be assisted in their online search. It rejects the proposition that consumers would necessarily rely on Google search as an entry point to obtain more shopping information.

(2) 'Thriving competition in online shopping'. Related to the previous point, Google argues that the market should not include only comparison shopping websites (as it is the case according to the Commission's definition), but rather it should be defined in a broader way so as to include shopping sites such as Amazon and eBay. More generally, Google claims that while a handful of price comparison aggregators may have lost clicks from Google search pages, there is a lot of competition in this (broader) market, which is dominated by Amazon. Data made publicly available by

¹⁵⁹ Conceivably, the same sort of theory of harm may also apply to other 'vertical shopping' services, such as flight, hotel, restaurants, that the Commission has been investigating before focusing on the comparison shopping market. In her press conference of 27 June 2017, the Competition Commissioner said that the comparison shopping decision may represent a precedent for similar abuses in other vertical markets.

¹⁶⁰ See for example the following entries in Google's blogs: <https://googleblog.blogspot.be/2015/04/the-search-for-harm.html?m=1>; <https://europe.googleblog.com/2015/08/improving-quality-isnt-anti-competitive.html>; <https://blog.google/topics/google-europe/improving-quality-isnt-anti-competitive-part-ii/>; www.blog.google/topics/google-europe/european-commission-decision-shopping-google-story/.

Google and which shows the evolution of monthly visitors of shopping sites in the UK from the end of 2006 until the end of 2014 seems to support Google's claims, in relation to the UK.¹⁶¹

However, if one only considered *comparison* shopping websites, one would obtain a very different view of the evolution of the market. Specifically, based on data publicly shared by one of the complainants in this case,¹⁶² there has been a general downward trend in monthly visits to UK comparison shopping websites between 2008 and 2014 (a period during which there was a very substantial increase in internet use and of e-commerce). Still, it may well be that this trend was not caused by Google's conduct, but by other factors, such as the rising importance of other services that facilitate online shopping.¹⁶³

The same complainant publicly shared data which also included the evolution of traffic of Google's own comparison shopping services, and extended the period over which Google's own services were included (relative to the data publicly shared by Google and referred to above). According to this data, the introduction of the change in the Google search algorithm (resulting in the 'Panda' version) in April 2011 appears to have been instrumental in promoting the sudden growth of Google's own service. The same data shows a fall in rivals' traffic since that date.¹⁶⁴

¹⁶¹ See in particular Google's blogs: <http://googleblog.blogspot.be/2015/04/the-search-for-harm.html?m=1>.

¹⁶² See in particular: www.foundem.co.uk/fmedia/Foundem_Jun_2015_Analysis/html5.html.

¹⁶³ It should also be noted that according to this data some rivals, such as Twenga and Idealprice, performed much better than others, and this begs the question of what is behind the different evolution. If it was because they have been able to innovate and offer better products, then Google's claim that comparison shopping sites have been failing because they have been unable to offer a better product would be vindicated. If it was instead because these sites have been much more aggressive in their use of online advertising, for instance because of reliance on Google AdWords, then it would be consistent with the complainants' arguments that Google stopped ranking them high in search pages in order to favour its own service or oblige rivals to rely on Google's advertising services.

¹⁶⁴ As the Commission's Press Release IP/17/1784 of 27 June 2017 also put it: 'Since the beginning of each abuse, Google's comparison shopping service has increased its traffic 45-fold in the United Kingdom, 35-fold in Germany, 19-fold in France, 29-fold in the Netherlands, 17-fold in Spain and 14-fold in Italy. Following the demotions applied by Google, traffic to rival comparison shopping services on the other hand dropped significantly. For example, the Commission found specific evidence of sudden drops of traffic to certain rival websites of 85 per cent in the United Kingdom, up to 92 per cent in Germany and 80 per cent in France. These sudden drops could also not be explained by other factors. Some competitors have adapted and managed to recover some traffic but never in full.'

The Commission argues that the evidence of these sudden changes in traffic demonstrates the simultaneous demotion of rivals and promotion of Google's services, and shows that 'Google's practices have stifled competition on the merits in comparison shopping markets, depriving European consumers of genuine choice and innovation'.¹⁶⁵

At the heart of the case there seems therefore to be a different idea of how consumers reach their online shopping decisions, which results into different market definitions and different effects on competition. In the Commission's view, consumers start looking for products through a search engine, then click through to a shopping comparison site, and from there they click on merchant websites. In Google's view, consumers reach merchant websites not only via general search engines, but also via specialist search services, merchant platforms, social networks, and online ads.

This is likely to be one of the thorniest issues that the judges will have to decide on, should this decision be appealed.

We limit ourselves to three considerations. First, the Commission's decision covers a long period, and it is very likely that the opportunities available to consumers may have expanded over time; and that while in 2008 they may have tended to start their online shopping by going to a search engine page, this behaviour may be much less pronounced ten years later (perhaps also because they have started to become loyal to particular sites over time).

Second, it is possible that there are at least two different groups of consumers, those who start their online shopping through general search and those, perhaps more sophisticated or simply more accustomed to online shopping, who go directly to merchant platforms or to merchant sites. Similarly, it is also possible that there are different types of merchants: those who decide to sell through merchant platforms like Amazon and eBay and those who prefer avoiding them – and who therefore would be more likely to be found starting from a general search page. If this was the case, then the Commission's reasoning may still hold, although it would not concern all consumers and all merchants but only particular groups of them.

Third, and related to the previous points, the appropriate market definition and the resulting effects on competition produced by Google's conduct may vary across countries. For instance, the data mentioned

¹⁶⁵ See the Commission's Press Release IP/17/1784 of 27 June 2017.

above referred to the evolution of traffic in the UK, where Amazon and eBay gained significant importance early on during the period of the investigation. Other countries, where online shopping has expanded later than in the UK, may display a different pattern.

Effects on consumers Another crucial point is whether the fall in rivals' traffic and the growth in Google's own services corresponded to the introduction of a drastic innovation in Google's comparison services. If that was the case, it is likely that consumers benefited from the innovation and the deterioration of rivals' position in the market would not be proof of anti-competitive harm.

Google has argued that they have provided consumers who intend to shop online with an improved service over time, in at least two ways.¹⁶⁶ First, by improving the format of their ads to include more informative displays with pictures, prices, and links where one can buy products. Second, by connecting users directly to merchants who sell the items, whether through organic links or ads. Google's claim is that its changes have improved consumer experience and that some price comparison aggregators have lost relevance because they were unable to provide the service that consumers wanted.

Again, it is worth noting that the abusive conduct at issue covers a long period, and that whatever innovative component may be in the service provided by Google, it is more likely to exist in the final years of the period, and in particular when (in 2013) 'Google shopping' introduced a 'box' containing pictures with price and product information (an auction decides which merchants would be selected in the box) which would bring consumers directly to the merchant sites. At the beginning of the period, instead, consumers would not be able to reach merchant sites directly from the Google search page, but would just go to the Froogle and Google Product Search sites, whose features were therefore more typical of a price comparison site.

More importantly, based on the evidence available in the public domain at the time of writing and summarised above, it is hard to see in the changes introduced in the search algorithm in April 2011 an innovative component that may benefit users (at least as far as comparison shopping services are concerned). Indeed, with reference for example to the UK case

¹⁶⁶ See, for example, www.blog.google/topics/google-europe/european-commission-decision-shopping-google-story (post of 27 June 2017).

discussed above, if the main driver of the fall in rivals' traffic was the change in Google's search algorithm, with an effect of steering traffic away from competitors and towards Google's own (and similar) services, then it would be plausible to conclude that consumer harm has occurred, by making it more difficult for consumers to reach a comparison shopping site they would have otherwise preferred.

References

- Abito, J.M. and J.D. Wright. 2008. "Exclusive Dealing with Imperfect Downstream Competition." *International Journal of Industrial Organization*. 26: 227–6.
- Adams, W.J. and J.L. Yellen. 1976. "Commodity Bundling and the Burden of Monopoly." *The Quarterly Journal of Economics*. 90: 475–98.
- Aghion, P. and P. Bolton. 1987. "Contracts as a Barrier to Entry." *American Economic Review*. 77: 388–401.
- Aguirre, I., S. Cowan and J. Vickers, 2010. "Monopoly Price Discrimination and Demand Curvature." *American Economic Review*. 100: 1601–15.
- Ahlborn, C. and D.S. Evans. 2009. "The Microsoft Judgment and Its Implications for Competition Policy Towards Dominant Firms in Europe." *Antitrust Law Journal*. 75: 887–932.
- Allain, M.-L., C. Chambolle and P. Rey. 2016. "Vertical Integration as a Source of Hold-up." *The Review of Economic Studies*. 83: 1–25.
- Almeida, H. and M. Campello. 2007. "Financial Constraints, Asset Tangibility and Corporate Investment." *The Review of Financial Studies*. 20: 1429–60.
- Almeida, H., C.S. Kim and H. Kim. 2016. "Internal Capital Markets in Business Groups: Evidence from the Asian Financial Crisis." *The Journal of Finance*. 70: 2539–86.
- Amato, G. 1997. *Antitrust and the Bounds of Power: The Dilemma of Liberal Democracy in the History of the Market*. Oxford: Hart Publishing.
- Amelio, A. and B. Jullien. 2012. "Tying and Freebies in Two-Sided Markets." *International Journal of Industrial Organization*. 30: 436–46.
- Andrews, P. 1998. "Aftermarket Power in the Computer Services Market: The Digital Undertaking." *European Competition Law Review*. 19: 176–81.
- Areeda, P.E. 1990. "Essential Facilities: An Epithet in Need of Limiting Principles." *Antitrust Law Journal*. 58: 841–54.
- Areeda, P.E. and H. Hovenkamp. 2003. *Antitrust Law: An Analysis of Antitrust Principles and Their Application*. Vol. IX, 2nd edn. New York: Aspen Publishers.
- Areeda, P.E. and H. Hovenkamp. 2006. *Antitrust Law: An Analysis of Antitrust Principles and Their Application*. New York: Aspen Publishers.

- Aareeda, P.E. and D.F. Turner. 1975. "Predatory Pricing and Related Practices under Section 2 of the Sherman Act." *Harvard Law Review*. 88: 697–733.
- Armstrong, M. 1996. "Multiproduct Nonlinear Pricing." *Econometrica*. 64: 51–75.
- Armstrong, M. 1998. "Network Interconnection in Telecommunications." *The Economic Journal*. 108: 545–64.
- Armstrong, M. 2006a. "Competition in two-sided markets." *RAND Journal of Economics*. 37: 668–91.
- Armstrong, M. 2006b. "Recent Developments in the Economics of Price Discrimination." In: Blundell, R., W. Newey and T. Persson, eds. *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, Vol. II*. Cambridge, UK: Cambridge University Press.
- Armstrong, M. 2008. "Price Discrimination." In: P. Buccirossi, ed. *Handbook of Antitrust Economics*. Cambridge, MA: MIT Press.
- Armstrong, M. and J. Vickers. 1993. "Price Discrimination, Competition and Regulation." *The Journal of Industrial Economics*. 41: 335–59.
- Armstrong, M. and J.D. Wright. 2007. "Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts." *Economic Theory*. 32: 107–12.
- Asker, J. 2016. "Diagnosing Foreclosure Due to Exclusive Dealing." *The Journal of Industrial Economics*. 64: 375–410.
- Asker, J. and H. Bar-Isaac. 2014. "Raising Retailers' Profits: On Vertical Practices and the Exclusion of Rivals." *American Economic Review*. 104: 672–86.
- Autorité de la Concurrence and Bundeskartellamt. 2016. "Competition Law and Data."
- Bacache, M., M. Bourreau and G. Gaudin. 2014. "Dynamic Entry and Investment in New Infrastructures: Empirical Evidence from the Fixed Broadband Industry." *Review of Industrial Organization*. 44: 179–209.
- Bakos, Y. and E. Brynjolfsson. 1999. "Bundling Information Goods: Pricing, Profits and Efficiency." *Management Science*. 45: 1613–30.
- Bartelsman, E.J. and M. Doms. 2000. "Understanding Productivity: Lessons from Longitudinal Microdata." *Journal of Economic Literature*. 38: 569–94.
- Baumol, W.J. 1996. "Predation and the Logic of the Average Variable Cost Test." *The Journal of Law and Economics*. 39: 49–72.
- Bernheim, B.D., B. Peleg and M. Whinston. 1987. "Coalition Proof Nash Equilibria I: Concepts." *Journal of Economic Theory*. 42: 1–12.
- Bernheim, B.D. and M.D. Whinston. 1998. "Exclusive Dealing." *Journal of Political Economy*. 106: 64–103.
- Besanko, D., U. Doraszelski and Y. Kryukov. 2014. "The Economics of Predation: What Drives Pricing When There Is Learning-By-Doing?" *American Economic Review*. 104: 868–97.
- Besanko, D. and M.K. Perry. 1993. "Equilibrium Incentives for Exclusive Dealing in a Differentiated Products Oligopoly." *RAND Journal of Economics*. 24: 646–67.
- Bhatia, V., A.I. Gavil, J.M. Jacobson, V. Lewis and J.B. McDonald. 2006. "Resources for Exclusive Dealing Agreements: A New Frontier in Antitrust Litigation?" The American Bar Association, Sherman Act Section 2 Committee.
- Bolton, P. and D. Scharfstein. 1990. "A Theory of Predation Based on Agency Problems in Financial Contracting." *American Economic Review*. 80: 93–106.
- Bolton, P., J. Brodley and M. Riordan. 2000. "Predatory Pricing: Strategic Theory and Legal Policy." *Georgetown Law Journal*. 88: 2239–330.

- Boone, J., W. Muller and S. Suetens. 2014. "Naked Exclusion in the Lab: The Case of Sequential Contracting." *The Journal of Industrial Economics*. 62: 137–66.
- Borenstein, S., J.K. MacKie-Mason and J.S. Netz. 1995. "Antitrust Policy in Aftermarkets." *Antitrust Law Journal*. 63: 455–82.
- Borenstein, S., J.K. MacKie-Mason and J.S. Netz. 2000. "Exercising Market Power in Proprietary Aftermarkets." *Journal of Economics and Management Strategy*. 9: 157–88.
- Bork, R. 1978. *The Antitrust Paradox: A Policy at War with Itself*. New York: Basic Books.
- Bouckaert, J. and F. Verboven. 2004. "Price Squeezes in a Regulatory Environment." *Journal of Regulatory Economics*. 26: 321–51.
- Boutin X., G. Cestone, C. Fumagalli, G. Pica and N. Serrano Velarde. 2013. "The Deep-Pocket Effect of Internal Capital Markets." *Journal of Financial Economics*. 109(1): 122–45.
- Bowman, W.S. 1957. "Tying Arrangements and the Leverage Problem." *The Yale Law Journal*. 67: 19–36.
- Braun, M. and B. Larrain. 2005. "Finance and the Business Cycle: International, Inter-Industry Evidence." *The Journal of Finance*. 60: 1097–1128.
- Broomhall, D., S.M. Nagel and U. Scholz. 2012. "Restrictive Practices in the Energy Sector: An Overview of EU and National Case Law." In Charbit, N., E. Ramundo and M. Courtaigne-Op, eds. *2013 Competition Law Digest – A synthesis of EU and national leading cases*. Institute of Competition Law.
- Burns, M.R. 1986. "Predatory Pricing and the Acquisition Cost of Competitors." *Journal of Political Economy*. 94: 266–96.
- Cabral, L.M.B. and M.H. Riordan. 1994. "The Learning Curve, Market Dominance, and Predatory Pricing." *Econometrica*. 62: 1115–40.
- Cabral, L.M.B. and M.H. Riordan. 1997. "The Learning Curve, Predation, Antitrust, and Welfare." *The Journal of Industrial Economics*. 45: 155–69.
- Cabral, L.M.B. 2014. "Staggered Contracts, Market Power, and Welfare." *CEPR Discussion Paper No. 10095*.
- Calcagno, C. and L. Giardino-Karlinger. 2016. "Collective Exclusion." *Mimeo*.
- Calzada, J. and T. Valletti. 2008. "Network Competition and Entry Deterrence." *The Economic Journal*. 118: 1223–44.
- Calzolari, G. and V. Denicolò. 2013. "Competition with Exclusive Contracts and Market-Share Discounts." *American Economic Review*. 103: 2384–411.
- Calzolari, G. and V. Denicolò. 2015. "Exclusive Contracts and Market Dominance." *American Economic Review*. 105: 3321–51.
- Calzolari, G., V. Denicolò and Zanchettin. 2016. "Exclusive Dealing with Costly Rent Extraction." *CEPR Discussion Paper No. 11291*.
- Carbajo, J., D. de Meza and J. Seidmann. 1990. "A Strategic Motivation for Commodity Bundling." *The Journal of Industrial Economics*. 38: 283–98.
- Cardoso, R., S. Kijewski, O. Koch, P. Lindberg and K. Nagy. 2010. "The Commission's GDF and E.ON Gas Decisions Concerning Long-Term Capacity Bookings – Use of Own Infrastructure as Possible Abuse under Article 102 TFEU." *Competition Policy Newsletter*. 2010–3: 8–11.
- Carlton, D.W. 2001. "A General Analysis of Exclusionary Conduct and Refusal to Deal – Why *Aspen* and *Kodak* are Misguided." *Antitrust Law Journal*. 68: 659–84.
- Carlton, D.W. 2008. "Should 'price squeeze' be a Recognised Form of Anticompetitive Conduct?" *Journal of Competition Law & Economics*. 4: 271–8.

- Carlton, D.W. and M. Waldman. 2002. "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries." *RAND Journal of Economics*. 33: 194–220.
- Carlton, D.W. and M. Waldman. 2010. "Competition, Monopoly, and Aftermarkets." *The Journal of Law, Economics, & Organization*. 26: 54–91.
- Carlton, D.W. and M. Waldman. 2012. "Upgrades, Switching Costs and the Leverage Theory of Tying." *Economic Journal*. 122: 675–726.
- Cave, M. 2006. "Encouraging Infrastructure Competition via the Ladder of Investment." *Telecommunications Policy*, 30: 223–37.
- Chapsal, A. 2014. "Repeated Purchases, Loyalty Rebates, and Demand Foreclosure." *Unpublished PhD thesis chapter*.
- Chen, C.-W. 2014. "Estimating the Foreclosure Effect of Exclusive Dealing: Evidence from the Entry of Specialty Beer Producers." *International Journal of Industrial Organization*. 37: 47–64.
- Chen, Y. 1997. "Equilibrium Product Bundling." *Journal of Business*. 70: 85–103.
- Chen, Z. and G. Shaffer. 2014. "Naked Exclusion with Minimum-Share Requirements." *RAND Journal of Economics*. 45: 64–91.
- Chen, C.-W. and S. Shieh. 2016. "Does Exclusive Dealing Matter? Evidence from Distribution Contract Changes in the US Beer Industry." *The Journal of Industrial Economics*. 64: 411–35.
- Choi, J.P. 1996. "Preemptive R&D, Rent Dissipation, and the Leverage Theory." *Quarterly Journal of Economics*. 110: 1153–81.
- Choi, J.P. 2004. "Tying and Innovation: A Dynamic Analysis of Tying Arrangements." *Economic Journal*. 114: 83–101.
- Choi, J.P. 2010. "Tying in Two-Sided Markets with Multi-Homing." *The Journal of Industrial Economics*. 58: 607–26.
- Choi, J.P. and C. Stefanadis. 2001. "Tying, Investment, and the Dynamic Leverage Theory." *RAND Journal of Economics*. 32: 52–71.
- Choi, J.P. and S.-S. Yi. 2000. "Vertical Foreclosure with the Choice of Input Specifications." *RAND Journal of Economics*. 31: 717–43.
- Choné, P. and L. Linnemer. 2015. "Nonlinear Ricing and Exclusion: I. Buyer Opportunism." *RAND Journal of Economics*. 46: 217–40.
- Church, J. and N. Gandal. 2000. "Systems Competition, Vertical Merger, and Foreclosure." *Journal of Economics & Management Strategy*. 9: 25–51.
- Competition and Markets Authority. 2015. "The Commercial Use of Consumer Data – Report on the CMA's call for information." *CMA 38*.
- Coppi, L. 2007. "Aftermarket Monopolization: The Emerging Consensus in Economics." *Antitrust Bulletin*. 52: 53–71.
- Corts, K.S. 1998. "Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment." *RAND Journal of Economics*. 29: 306–23.
- Cowan, S. 2012. "Third-Degree Price Discrimination and Consumer Surplus." *The Journal of Industrial Economics*. 60: 333–45.
- Crane, D.A. 2005. "The Paradox of Predatory Pricing." *Cornell Law Review*. 91: 1–66.
- Crane, D.A., F.P. Furth, K.G. Elzinga, R.A. Epstein, J.A. Hausman, R. Haw, H. Hovenkamp, R.G. Hubbard, K. Hylton, W.E. Kovacic, T. Lambert, A. Meese, T.D. Morgan, B.Y. Orbach, W.H. Page, R.S. Pindyck, E.A. Snyder, D.D. Sokol and R.H. Topel. 2013. "Brief for eighteen scholars as Amici Curiae in support of petitioner."

- On petition for writ of certiorari to the U.S. Court of Appeals for the Third Circuit, in *Eaton Corp. v ZF Meritor LLC and Meritor Transmission Corp* (No. 12–1045). 28 March 2013.
- Crawford, G.S. and A. Yurukoglu. 2012. “The Welfare Effects of Bundling in Multichannel Television Markets.” *American Economic Review*. 102: 181–219.
- Cummings, F.J. and W.E. Ruhter. 1979. “The Northern Pacific Case.” *The Journal of Law and Economics*. 22: 329–50.
- Dalton, J.A. and L. Esposito. 2007. “Predatory Price Cutting and Standard Oil: A Re-examination of the Trial Record.” In Zerbe, R. and J. Kirkwood, eds. *Research in Law and Economics*. Vol. 22: 155–205. Emerald Group Publishing Limited.
- Dana, J.D. and K.E. Spier. 2015. “Bundling and quality assurance.” *Mimeo*.
- De Meza, D. and M. Selvaggi. 2007. “Exclusive Contracts Foster Relationship-Specific Investment.” *RAND Journal of Economics*. 38: 85–97.
- DeGraba, P. 1990. “Input Market Price Discrimination and the Choice of Technology.” *American Economic Review*. 80: 1246–53.
- Dewatripont, M. and E. Maskin. 1995. “Credit and Efficiency in Centralized and Decentralized Economies.” *The Review of Economic Studies*. 62: 541–55.
- Directorate-General for Competition (European Commission). 2005. “DG Competition discussion paper on the application of Article 82 of the Treaty to exclusionary abuses.”
- Doganoglu, T. and J.D. Wright. 2010. “Exclusive Dealing with Network Effects.” *International Journal of Industrial Organization*. 28: 145–54.
- Dessein, W. 2003. “Network Competition in Nonlinear Pricing.” *RAND Journal of Economics*. 34: 1–19.
- Economides, N. and I. Lianos. 2009. “The elusive antitrust standard on bundling in Europe and in the United States in the aftermath of the Microsoft cases.” *Antitrust Law Journal*. 76: 483–567.
- Economides, N. and I. Lianos. 2010. “A Critical Appraisal of Remedies in the E.U. Microsoft Cases.” *Columbia Business Law Review*. 2010: 346–420.
- Edlin, A.S. and J. Farrell. 2004. “The American Airlines Case: A Chance to Clarify Predation Policy (2001).” In Kwoka, J.E. and L.J. White, eds. *The Antitrust Revolution: Economics, Competition, and Policy*. 4th ed. Oxford: Oxford University Press.
- Edwards, G. and M. Walker. 2013. “When is a Margin Squeeze Not an Abuse?” *CRA Competition Memo*. July 2013.
- Elhauge, E. 2003. “Why Above-Cost Price Cuts To Drive Out Entrants Are Not Predatory – and the Implications for Defining Costs and Market Power.” *The Yale Law Journal*. 112: 681–827.
- Elhauge, E. 2009. “How Loyalty Discounts May Perversely Discourage Discounting.” *Journal of Competition Law & Economics*. 5: 189–231.
- Elzinga, K.G. 1989. “Collusive Predation.” In Kwoka, J.E. and L.J. White, eds. *The Antitrust Revolution*. 1st edn. Glenview, IL: Scott, Foresman and Company.
- Elzinga, K.G. and D.E. Mills. 2001. “Independent Service Organizations and Economic Efficiency.” *Economic Inquiry*. 39: 549–60.
- European Commission. 2009. “Communication from the Commission Guidance on the Commission’s Enforcement Priorities in Applying Article 102 TFEU to Abusive Exclusionary Conduct by Dominant Undertakings.” OJ [2009] C 45/07.

- European Commission. 2010. "Commission Notice: Guidelines on Vertical Restraints." OJ [2010] C130/01.
- European Regulators Group. 2008. "ERG Report on the Consultation for the ERG Common Position on Symmetry of Fixed Call Termination Rates and Symmetry of Mobile Call Termination Rates." Adopted on 28 February 2008 (ERG (07) 83 final 080312).
- Evans, H. 1912. "The Standard Oil and American Tobacco Cases." *University of Pennsylvania Law Review and American Law Register*. 60: 311–23.
- Evans, D.S. 2003. "The Antitrust Economics of Multi-Sided Platform Markets." *Yale Journal on Regulation*. 20: 325–81.
- Evans, D.S. 2009. "Why Different Jurisdictions Do Not (and Should Not) Adopt the Same Antitrust Rules." *The Chicago Journal of International Law*. 10: 161–87.
- Evans, D.S. and M.A. Salinger. 2005. "Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law," *Yale Journal on Regulation*. 22: 37–89.
- Evans, D.S. and M.A. Salinger. 2007. "Curing Sinus Headaches and Tying Law," in Choi, J.P., ed., *Recent Developments in Antitrust*. Cambridge, MA: MIT Press.
- Ezrachi, A. 2016. *EU Competition Law: An Analytical Guide to the Leading Cases*. 5th edn. Oxford: Hart Publishing.
- Fang, H. and P. Norman. 2006. "To Bundle or Not to Bundle." *RAND Journal of Economics*. 37: 946–63.
- Farrell, J. 2005. "Deconstructing Chicago on Exclusive Dealing." *The Antitrust Bulletin*. 50: 465–80.
- Farrell, J. and M.L. Katz. 2005. "Competition or Predation? Consumer Coordination, Strategic Pricing and Price Floors in Network Markets." *The Journal of Industrial Economics*. 53: 203–31.
- Fazzari, S.R., G. Hubbard and B. Petersen. 1988. "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity*. 1: 141–95.
- FCA. 2013. "Applying Behavioural Economics at the Financial Conduct Authority." *Occasional Paper No. 1*.
- Federico, G. 2011a. "The Antitrust Treatment of Loyalty Discounts in Europe: Towards a more Economic Approach." *Journal of European Competition Law & Practice*. 2: 277–84.
- Federico, G. 2011b. "Tomra v. Commission of the European Communities: Reversing Progress on Rebates?" *European Competition Law Review*. 32: 139–41.
- Federico, G. 2013. "SAA II: Abuse of Dominance in the South African Skies." *Journal of Competition Law & Economics*. 9: 709–37.
- Federico, G. and P. Régibeau. 2012. "Abusive Incentive Contracts: British Airways Revisited." *Mimeo*.
- Federico, G. and X. Vives. 2008. "Competition and Regulation in the Spanish Gas and Electricity Markets." *Reports of the Public-Private Sector Research Center (IESE Business School)*, 1.
- Fox, E.M. 2005. "Is There Life in Aspen after *Trinko*? The Silent Revolution of Section 2 of the Sherman Act." *Antitrust Law Journal*. 73: 153–70.
- Friederiszick, H. W. and L. Gratz. 2012. "Dominant and Efficient: On the Relevance of Efficiencies in Abuse of Dominance Cases." *ESMT White Paper No. WP-12-01*.

- Fudenberg, D. and J. Tirole. 1984. "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look." *American Economic Review*. 74: 361–66.
- Fudenberg, D. and J. Tirole. 1986. "A Signal-Jamming Theory of Predation." *RAND Journal of Economics*. 17: 173–90.
- Fumagalli, C. and M. Motta. 2006. "Exclusive Dealing and Entry, When Buyers Compete." *American Economic Review*. 96: 785–95.
- Fumagalli, C. and M. Motta. 2008. "Buyers' Miscoordination, Entry, and Downstream Competition." *The Economic Journal*. 118: 1196–222.
- Fumagalli, C. and M. Motta. 2013. "A Simple Theory of Predation." *The Journal of Law and Economics*. 56: 595–631.
- Fumagalli, C. and M. Motta. 2017a. "On the Use of Price-cost Tests in Loyalty Discounts and Exclusive Dealing Arrangements: Which Implications from Economic Theory Should Be Drawn?" *Antitrust Law Journal*. 81: 537–85.
- Fumagalli, C. and M. Motta. 2017b. "Dynamic Vertical Foreclosure." *Mimeo*.
- Fumagalli, C., M. Motta and L. Persson. 2009. "On the Anticompetitive Effect of Exclusive Dealing When Entry by Merger is Possible." *The Journal of Industrial Economics*. 47: 785–811.
- Fumagalli, C., M. Motta and T. Rønne. 2012. "Exclusive Dealing: Investment Promotion May Facilitate Inefficient Foreclosure." *The Journal of Industrial Economics*. 60: 599–608.
- Gabel, D. 1994. "Competition in a Network Industry: The Telephone Industry, 1894–1910." *The Journal of Economic History*. 54: 543–72.
- Gabel, D. and D.I. Rosenbaum. 1995. "Prices, Costs, Externalities and Entrepreneurial Capital: Lessons from Wisconsin." *The Antitrust Bulletin*. 40: 581–608.
- Gans, J.S. and S. King. 2001. "Using Bill and Keep Interconnection Arrangements to Soften Network Competition." *Economic Letters*. 71: 413–20.
- Gans, J.S. and S. King. 2002. "Exclusionary Contracts and Competition for Large Buyers." *International Journal of Industrial Organization*. 7: 381–401.
- Gavil, A.I., W.E. Kovacic and J.B. Baker. 2008. *Antitrust Law in Perspective: Cases, Concepts, and Problems in Competition Policy*. 2nd ed. St. Paul, MN: Thomson/West.
- Genakos, C. and T. Valletti. 2011. "Testing the Waterbed Effect in Mobile Telephony." *Journal of the European Economic Association*. 9: 1114–42.
- Geradin, D., A. Layne-Farrar and N. Petit. 2012. *EU Competition Law and Economics*. Oxford: Oxford University Press.
- Geradin, D. and R. O'Donoghue. 2005. "The Concurrent Application of Competition Law and Regulation: The Case of Margin Squeeze Abuses in the Telecommunications Sector." *Journal of Competition Law & Economics*. 1: 355–425.
- Gilbert, R.J. and M.L. Katz. 2001. "An Economist's Guide to *U.S. v. Microsoft*." *Journal of Economic Perspectives*. 15: 25–44.
- Goldfine, D.A.J. and K.M. Vorrasi. 2004. "The Fall of the Aftermarket Kodak Doctrine: Dying a Slow Death in the Lower Courts." *Antitrust Law Journal*. 72: 209–32.
- Greenlee, P., D. Reitman and D.S. Sibley. 2008. "An antitrust analysis of bundled loyalty discounts." *International Journal of Industrial Organization*. 26: 1132–52.
- Groh, C. and G. Spagnolo. 2004. "Exclusive Contracts, Loss to Delay and Incentives to Invest." *CEPR Discussion Paper No. 4525*.

- Hajivassiliou, V. and F. Savignac. 2016. "Financing Constraints and a Firm's Decision and Ability to Innovate: Novel Approaches to Coherency Conditions in Dynamic LDV Models." *Mimeo*.
- Hansen, S. and M. Motta. 2012. "Vertical Exclusion with Endogenous Competition Externalities." *CEPR Discussion Paper No. 8982*.
- Harrington, J.E. 1986. "Limit Pricing When the Potential Entrant Is Uncertain of Its Cost Function." *Econometrica*. 54: 429–37.
- Harrington, J.E. 1989. "Collusion and Predation Under (almost) Free Entry." *International Journal of Industrial Organization*. 7: 381–401.
- Hart, O. and J. Tirole. 1990. "Vertical Integration and Market Foreclosure." *Brookings Papers on Economic Activity (Microeconomics)*. 1990: 205–85.
- Hausman, J.A. 1979. "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *The Bell Journal of Economics*. 10: 33–54.
- Heide, J.B., S. Dutta and M. Bergen. 1998. "Exclusive Dealing and Business Efficiency: Evidence from Business Practice." *The Journal of Law and Economics*. 41: 387–407.
- Heimler, A. 2010. "Is Margin Squeeze an Antitrust or a Regulatory Violation?" *Journal of Competition Law & Economics*. 6: 879–90.
- Hilke, J.C. and P.B. Nelson. 1987. "Caveat Innovator – Strategic and Structural Characteristics of New Product Introductions." *Journal of Economic Behavior and Organization*. 8: 213–29.
- Hovenkamp, H. 2001. "Post-Chicago Antitrust: A Review and Critique." *Columbia Business Law Review*. 2001: 257–338.
- Hovenkamp, H. 2005. *Federal Antitrust Policy: The Law of Competition and Its Practice*. St. Paul, MN: Thomson/West.
- Hurkens, S., D.-S. Jeon and D. Menicucci. 2016. "Leveraging Dominance with Credible Bundling." *CEPR Discussion Paper No. 11304*.
- Hylton, K.N. 2003. *Antitrust Law: Economic Theory & Common Law Evolution*. Cambridge: Cambridge University Press.
- Ide, E., J.-P. Montero and N. Figueroa. 2016. "Discounts as a Barrier to Entry." *American Economic Review*. 106: 1849–77.
- Inderst, R. and N. Mazzarotto. 2007. "Buyer Power in Distribution." In Collins, W.D., ed., *Issues in Competition Law and Policy*. Chicago: American Bar Association.
- Inderst, R. and G. Shaffer. 2009. "Market Power, Price Discrimination, and Allocative Efficiency in Intermedate-Goods Markets." *RAND Journal of Economics*. 40: 658–72.
- Inderst, R. and G. Shaffer. 2010. "Market-Shares Contracts as Facilitating Practices." *RAND Journal of Economics*. 41: 709–29.
- Inderst, R. and T. Valletti. 2009. "Price Discrimination in Input Markets." *RAND Journal of Economics*. 40: 1–19.
- Innes, R. and R.J. Sexton. 1993. "Customer Coalitions, Monopoly Price Discrimination and Generic Entry Deterrence." *European Economic Review*. 37: 1569–97.
- Jacobson, J.M. 2002. "Exclusive Dealing, 'Foreclosure', and Consumer Harm." *Antitrust Law Journal*. 70: 311–69.
- Jeon, D.-S. and D. Menicucci. 2006. "Bundling Electronic Journals and Competition Among Publishers." *Journal of the European Economic Association*. 4: 1038–83.
- Jeon, D.-S. and D. Menicucci. 2012. "Bundling and Competition for Slots." *American Economic Review*. 102: 1957–85.

- Jing, R. and R.A. Winters. 2014. "Exclusionary Contracts." *The Journal of Law, Economics & Organization*. 30: 833–67.
- Johnson, J. 2012. "Adverse Selection and Partial Exclusive Dealing." *Mimeo*.
- Jones, A. and B. Sufirin. 2016. *EU Competition Law: Text, Cases and Materials*. Oxford: Oxford University Press.
- Joskow, P.L. and A.K. Klevorick. 1979. "A Framework for Analyzing Predatory Pricing Policy." *The Yale Law Journal*. 89: 213–70.
- Jullien, B., P. Rey and C. Saavedra. 2014. "The Economics of Margin Squeeze." *CEPR Discussion Paper No. 9905*.
- Karlinger, L. and M. Motta. 2012. "Exclusionary Pricing When Scale Matters." *The Journal of Industrial Economics*. 60: 75–103.
- Katz, M.L. and C. Shapiro. 1986. "Technology Adoption in the Presence of Network Externalities." *Journal of Political Economy*. 94: 822–41.
- Kim, S-H. 2009. "Predatory Reputation in US Airline Markets." *International Journal of Industrial Organization*. 27: 592–604.
- Klein, B. 1996. "Market Power in Aftermarkets." *Managerial and Decision Economics*. 17: 143–64.
- Kobayashi, B.H. 2005. "The Economics of Loyalty Discounts and Antitrust Law in the United States." *Competition Policy International*. 1: 115–48.
- Kobayashi, B.H. 2010. "The Law and Economics of Predatory Pricing." In Hylton, K. N., ed. *Antitrust Law And Economics*. Cheltenham and Camberley: Edward Elgar.
- Kolay, K., G. Shaffer and Ordovery, J.A. 2004. "All-Units Discounts in Retail Contracts." *Journal of Economics & Management Strategy*. 13: 429–59.
- Korah, V. 2006. *Cases and Materials on EC Competition Law*. Oxford: Hart.
- Korah, V. 2007. *An Introductory Guide to EC Competition Law and Practice*. Oxford: Hart.
- Kovacic, K.E. 2007. "The Intellectual DNA of Modern U.S. Competition Law for Dominant Firm Conduct: The Chicago/Harvard Double Helix." *Columbia Business Law Review*. 2007: 1–82.
- Krattenmaker, T.G. and S.C. Salop. 1986. "Anticompetitive Exclusion: Raising Rivals' Costs to Achieve Power over Price." *The Yale Law Journal*. 96: 209–93.
- Kreps, D.M. and J.A. Scheinkman. 1983. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics*. 14: 326–37.
- Kreps, D.M. and R. Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory*. 27: 253–79.
- Kühn, K.-U., R. Stillman and C. Caffarra. 2005. "Economic Theories of Bundling and Their Policy Implications in Abuse Cases: An Assessment in Light of the Microsoft Case." *European Competition Journal*. 1: 85–121.
- Kühn, K.-U., and J. Van Reenen. 2009. "Interoperability and Market Foreclosure in the European Microsoft Case." In Bruce Lyons, ed. *Cases in European Competition Policy*. Cambridge: Cambridge University Press.
- Kuppuswamy, V. and B. Villalonga. 2016. "Does Diversification Create Value in the Presence of External Financing Constraints? Evidence from the 2007–2009 Financial Crisis." *Management Science*. 62: 905–23.
- Laffont, J.J., P. Rey and J. Tirole. 1998a, "Network Competition: I. Overview and Nondiscriminatory Pricing." *RAND Journal of Economics*. 29: 1–37.

- Laffont, J.J., P. Rey and J. Tirole. 1998b. "Network Competition: II. Price Discrimination." *RAND Journal of Economics*. 29: 38–56.
- Landeo, C.M. and K.E. Spier. 2009. "Naked Exclusion: An Experimental Study of Contracts with Externalities." *American Economic Review*. 99: 1850–77.
- Larouche, P. 2008. "The European Microsoft Case at the Crossroads of Competition Policy and Innovation." *Antitrust Law Journal*. 75: 601–31.
- Lerner, J. 1995. "Pricing and Financial Resources: An Analysis of the Disk Drive Industry, 1980–88." *The Review of Economics and Statistics*. 77: 585–98.
- Lopez, A.L. and P. Rey. 2016. "Foreclosing Competition Through High Access Charges and Price Discrimination." *The Journal of Industrial Economics*. 64: 436–65.
- Luo, Y. 2013. "Bundling and Nonlinear Pricing in Telecommunications." *PhD thesis*.
- Maier-Rigaud, F., F. Manca and U. von Koppenfels. 2011. "Strategic Underinvestment and Gas Network Foreclosure: The ENI Case." *Competition Policy Newsletter*. 2011 (1): 18–23.
- Maier-Rigaud, F. and D. Vaigauskaite. 2006. "Prokent/Tomra, A Textbook Case? Abuse of Dominance under Perfect Information." *Competition Policy Newsletter*. 2006 (2): 19–24.
- Majudmar, A. and Shaffer G. 2009. "Market-Share Contracts with Asymmetric Information." *Journal of Economics & Management Strategy*. 18: 393–421.
- Maksimovic, V. and G. Phillips. 2008. "The Industry Life Cycle, Acquisitions and Investments: Does Firm Organization Matter?" *The Journal of Finance*. 63: 673–708.
- Mankiw, N.G. and M.D. Whinston. 1986. "Free Entry and Social Inefficiency." *RAND Journal of Economics*. 17: 48–58.
- Marin, P.L., and R. Sicotte. 2003. "Exclusive Contracts and Market Power: Evidence from Ocean Shipping." *The Journal of Industrial Economics*. 51: 193–213.
- Martimort, D. 1996. "Exclusive Dealing, Common Agency, and Multiprincipals Incentive Theory." *RAND Journal of Economics*. 27: 1–31.
- Marvel, H.P. 1982. "Exclusive Dealing." *The Journal of Law and Economics*. 25: 1–26.
- Marx, L.M. and G. Shaffer. 2007. "Upfront Payments and Exclusion in Downstream Markets." *RAND Journal of Economics*. 38: 823–43.
- Mathewson, G.F. and R.A. Winter. 1987. "The Competitive Effects of Vertical Agreements: Comment." *American Economic Review*. 77: 1057–62.
- Mathewson, F. and R. Winter. 1997. "Tying as a Response to Demand Uncertainty." *RAND Journal of Economics*. 28: 566–83.
- McAfee, R.P., J. McMillan and M. Whinston. 1989. "Multiproduct Monopoly, Commodity Bundling and Correlation of Values." *Quarterly Journal of Economics*. 104: 371–84.
- McAfee, R.P. and M. Schwartz. 1994. "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity." *American Economic Review*. 84: 210–30.
- McGee, J.S. 1958. "Predatory Price Cutting: The Standard Oil (N. J.) Case." *The Journal of Law and Economics*. 1: 137–69.
- Meese, A.J. 2005. "Property, *Aspen* and Refusals to Deal." *Antitrust Law Journal*. 73: 81–114.
- Miklos-Thal, J., P. Rey and T. Vergé. 2011. "Buyer Power and Intra-brand Coordination." *Journal of the European Economic Association*. 9: 721–41.

- Milgrom, P. and J. Roberts. 1982. "Predation, Reputation and Entry Deterrence." *Journal of Economic Theory*. 27: 280–312.
- Milliou, C. 2008. "Technological Proximity and Exclusive Buyer-Supplier Relationships." *The B.E. Journal of Economic Analysis & Policy*. 8(1).
- Mills, D. 2010. "Inducing Downstream Selling Effort through Market-Share Discounts." *International Journal of the Economics of Business*. 17: 129–46.
- Motta, M. 2004. *Competition Policy. Theory and Practice*. Cambridge: Cambridge University Press.
- Motta, M. 2009. "Michelin II – The Treatment of Rebates." In Lyons, B., ed. *Cases in European Competition Policy*. Cambridge, UK: Cambridge University Press.
- Motta, M. and A. de Streel. 2006. "Excessive Pricing and Price Squeeze under EU Law." In Ehlermann, D. and I. Atanasiu, eds. *What is an Abuse of a Dominant Position?* Oxford: Hart Publishing.
- Motta, M. and A. de Streel. 2007. "Excessive Pricing in Competition Law: Never Say Never?" in *The Pros and Cons of High Prices*. Stockholm: Konkurrensverket (Swedish Competition Authority).
- Nalebuff, B. 2004. "Bundling as an Entry Barrier." *Quarterly Journal of Economics*. 119: 159–88.
- Nalebuff, B. 2005. "Exclusionary Bundling." *The Antitrust Bulletin*. 50: 321–70.
- Nalebuff, B. and D. Majerus. 2003. "Bundling, Tying, and Portfolio Effects – Part 2 Case Studies." *DTI Economics Paper No. 1*.
- Nurski, L. and F. Verboven. 2016. "Exclusive Dealing as a Barrier to Entry? Evidence from Automobiles." *The Review of Economic Studies*. 83: 1156–88.
- O'Brien, D.P. and G. Shaffer. 1992. "Vertical Control with Bilateral Contracts." *RAND Journal of Economics*. 23: 299–308.
- O'Donoghue, R. and A.J. Padilla. 2013. *The Law and Economics of Article 102 TFEU*. Oxford: Hart Publishing.
- Ochs, J. 1995. "Coordination Problems." In: Kagel, J. H. and Roth, A. E., eds. *The Handbook of Experimental Economics*. NJ: Princeton University Press.
- OECD. 2009. Directorate for Financial and Enterprise Affairs, Competition Committee, Working Party No. 2 on Competition and Regulation. "Margin Squeeze" (contribution by the United States). *DAF/COMP/WP2/WD(2009)31*.
- OFT. 2005. "Selective Price Cuts and Fidelity Rebates." A report prepared for the Office of Fair Trading by RBB Economics. *Economic Discussion Paper (OFT 804)*. July 2005.
- Ordovery, J.A., G. Saloner and S.C. Salop. "Equilibrium Vertical Foreclosure." *American Economic Review*. 80: 127–42.
- Ordovery, J.A. and R.D. Willig. 1981. "An Economic Definition of Predation: Pricing and Product Innovation." *The Yale Law Journal*. 91: 8–53.
- Peitz, M. 2008. "Bundling May Blockade Entry." *International Journal of Industrial Organization*. 26: 41–58.
- Peterman, J.L. 1979. "The International Salt Case." *The Journal of Law and Economics*. 22: 351–64.
- Pitofsky, R., D. Patterson and J. Hooks. 2002. "The Essential Facilities Doctrine Under U.S. Antitrust Law." *Antitrust Law Journal*. 70: 443–62.
- Posner, R.A. 1976. *Antitrust Law: An Economic Perspective*. Chicago: University of Chicago Press.
- Posner, R.A. 2001. *Antitrust Law*. Chicago: University of Chicago Press.

- Rajan, R. and L. Zingales. 1998. "Financial Dependence and Growth." *American Economic Review*. 88: 559–86.
- Rasmusen, E.B., J.M. Ramseyer and J.J.S. Wiley. 1991. "Naked Exclusion." *American Economic Review*. 81: 1137–45.
- Reisinger, M. and E. Tarantino. 2015. "Vertical Integration, Foreclosure, and Productive Efficiency." *RAND Journal of Economics*. 46: 461–79.
- Rey, P. and J. Tirole. 1986. "The Logic of Vertical Restraints." *American Economic Review*. 76: 921–39.
- Rey, P. and J. Tirole. 2007. "A Primer on Foreclosure." In Armstrong, M. and R. Porter, eds. *Handbook of Industrial Organization (Vol. 3)*. Amsterdam: North-Holland.
- Rey, P. and T. Vergé. 2004. "Bilateral Control with Vertical Contracts." *RAND Journal of Economics*. 35: 728–46.
- Rey, P. and M.D. Whinston. 2013. "Does Retailer Power Lead to Exclusion?" *RAND Journal of Economics*. 44: 75–81.
- Ridyard, D. 1996. "Essential Facilities and the Obligation to Supply competitors under the UK and EC Competition Law." *European Competition Law Review*. 17: 438–52.
- Rochet, J.C. and J. Tirole. 2003. "Platform Competition in Two-Sided Markets." *Journal of the European Economic Association*. 1: 990–1029.
- Rochet, J.C. and J. Tirole. 2006. "Two-Sided Markets: A Progress Report." *RAND Journal of Economics*. 37: 645–67.
- Rochet, J.C. and J. Tirole. 2008. "Tying in Two-Sided Markets and the Honor all Cards Rule." *International Journal of Industrial Organization*. 26: 1333–47.
- Rubinfeld, D.L. 2005. "3M's Bundled Rebates: An Economic Perspective." *The University of Chicago Law Review*. 72: 243–64.
- Russo, F., M.P. Schinkel, A. Günster and M. Carree. 2010. *European Commission Decisions on Competition: Economic Perspectives on Landmark Antitrust and Merger Cases*. Cambridge: Cambridge University Press.
- Sagers, C. 2009. "Rarely tried, and... Rarely Successful: Theoretically Impossible Price Predation among the Airlines." *Journal of Air Law and Commerce*. 74: 919–58.
- Salinger, M.A. 1995. "A Graphical Analysis of Bundling." *The Journal of Business*. 68: 85–98.
- Saloner, G. 1987. "Predation, Mergers, and Incomplete Information." *RAND Journal of Economics*. 18: 165–86.
- Salop, S.C. 2006. "Exclusionary Conduct, Effect on Consumers, and the Flawed Profit-Sacrifice standard." *Antitrust Law Journal*. 73: 311–74.
- Salop, S.C. 2010. "Refusals to Deal and Price Squeezes by an Unregulated, Vertically Integrated Monopolist." *Antitrust Law Journal*. 76: 709–32.
- Salop, S.C. and D.T. Scheffman. 1983. "Raising Rivals' Costs." *American Economic Review (Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association)*. 73: 267–71.
- Salop, S.C. and D.T. Scheffman. 1987. "Cost-Raising Strategies." *The Journal of Industrial Economics*. 36: 19–34.
- Sappington, D.E.M. and J.G. Sidak. 2003. "Competition Law for State-Owned Enterprises." *Antitrust Law Journal*. 71: 479–524.
- Sass, T.R. 2005. "The Competitive Effects of Exclusive Dealing: Evidence from the U.S. Beer Industry." *International Journal of Industrial Organization*. 23: 203–25

- Sass, T.R. and M. Gisser. 1989. "Agency Cost, Firm Size, and Exclusive Dealing." *The Journal of Law and Economics*. 32: 381–400.
- Scharfstein, D.S. 1984. "A Policy to Prevent Rational Test–Marketing Predation." *RAND Journal of Economics*. 2: 229–43.
- Schmalensee, R. 1981. "Output and welfare effects of monopolistic third-degree price discrimination." *American Economic Review*. 71: 242–7.
- Schmalensee, R. 1982. "Commodity Bundling by Single-Product Monopolies." *The Journal of Law and Economics*. 25: 67–71.
- Schmalensee, R. 1984. "Gaussian Demand and Commodity Bundling." *The Journal of Business*. 57: 211–30.
- Schmalensee, R. 2002. "Payment Systems and Interchange Fees." *The Journal of Industrial Economics*. 50: 103–22.
- Schwartz, M. and G.J. Werden. 1996. "A Quality-Signaling Rationale for Aftermarket Tying." *Antitrust Law Journal*. 64: 387–404.
- Scott Morton, F. 1997. "Entry and Predation: British Shipping Cartels 1879–1929." *Journal of Economics & Management Strategy*. 6: 679–724.
- Segal, I.R. 1999. "Contracting with Externalities." *The Quarterly Journal of Economics*. 114: 337–88.
- Segal, I.R. 2003. "Coordination and Discrimination in Contracting with Externalities: Divide and Conquer?" *Journal of Economic Theory*. 113: 147–81.
- Segal, I.R. and M.D. Whinston. 2000a. "Naked Exclusion: Comment." *American Economic Review*. 90: 296–309.
- Segal, I.R. and M.D. Whinston. 2000b. "Exclusive Contracts and Protection of Investments." *RAND Journal of Economics*. 31: 603–33.
- Segal, I.R. and M.D. Whinston. 2007. "Antitrust in Innovative Industries." *American Economic Review*. 97: 1703–30.
- Selten, R. 1975. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *International Journal of Game Theory*. 4: 25–55.
- Shapiro, C. 1995. "Aftermarkets and Consumer Welfare: Making Sense of Kodak." *Antitrust Law Journal*. 63: 483–511.
- Shelanski, H.A. 2009. "Unilateral Refusals to Deal in Intellectual and Other Property." *Antitrust Law Journal*. 76: 369–95.
- Sher, S.A. and S.D. Russell. 2005. "Adding Bite to Exclusive Dealing? An Analysis of the Third Circuit's Dentsply Decision." *The Antitrust Source*. May 2005.
- Simpson, J. and A.L. Wickelgren. 2007. "Naked Exclusion, Efficient Breach, and Downstream Competition." *American Economic Review*. 97: 1305–20.
- Smith, A.M. 2011. "An Experimental Study of Exclusive Contracts." *International Journal of Industrial Organization*. 29: 4–13.
- Spector, D. 2011. "Exclusive Contracts and Demand Foreclosure." *RAND Journal of Economics*. 42: 619–38.
- Spier, K.E. and M.D. Whinston. 1995. "On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance and Renegotiation." *RAND Journal of Economics*. 26: 180–202.
- Stole, L.A. 2007. "Price discrimination and competition." In Armstrong, M. and R. Porter, eds. *Handbook of Industrial Organization (Vol. 3)*. Amsterdam: North-Holland.

- Syverson, C. 2011. "What Determines Productivity?" *Journal of Economic Literature*. 49: 326–65.
- Tapia, J. and S. Roberts. 2015. "Abuses of Dominance in Developing Countries: A View from the South, with an Eye on Telecommunications." In Gal, M.S., M. Bakhoun, J. Drexler, E.M. Fox, and D.J. Gerber, eds. *The Economic Characteristics of Developing Jurisdictions – Their Implications for Competition Law*. Cheltenham, UK: Edward Elgar.
- Telser, L.G. 1966. "Cutthroat Competition and the Long Purse." *The Journal of Law and Economics*. 9: 259–77.
- Temple Lang, J. 2011. "Practical Aspects of Aftermarkets in European Competition Law." *Competition Policy International*. 7: 199–243.
- Thisse J.-F. and X. Vives. 1988. "On the Strategic Choice of Spatial Price Policy." *American Economic Review*. 78: 122–37.
- Tirole, J. 1988. *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.
- Tirole, J. 2005. "The Analysis of Tying Cases: A Primer." *Competition Policy International*. 1: 1–25.
- Varian, H. 1989. "Price Discrimination." In Schmalensee, R. and R.D. Willig, eds. *Handbook of Industrial Organization (Vol. 1)*. Amsterdam: North-Holland.
- Vasconcelos, L. 2014. "Contractual Signaling, Relationship-Specific Investment and Exclusive Agreements." *Games and Economic Behavior*. 87: 19–33.
- Veljanovski, C. 2013. "Margin Squeeze and Competition Law: An Overview of EU and National Case Law." In Charbit, N., E. Ramundo and M. Op-Courtaigne, eds. *2013 Competition Law Digest – A synthesis of EU and national leading cases*. Institute of Competition Law.
- Vickers, J. 2005. "Abuse of Market Power." *The Economic Journal*. 115: F244–F261.
- Vickers, J. 2007. "Competition Law and Economics: A Mid-Atlantic Viewpoint." *European Competition Journal*. 3: 1–15.
- Von Rosenberg, H. 2009. "Unbundling through the Back Door...the Case of Network Divestiture as a Remedy in the Energy Sector." *European Competition Law Review*. 30: 237–54.
- Waller, S.W. and W. Tasch. 2010. "Harmonizing Essential Facilities." *Antitrust Law Journal*. 76: 741–62.
- Weiman, D.F. and R.C. Levin. 1994. "Preying for Monopoly? The Case of Southern Bell Telephone Company, 1894–1912." *Journal of Political Economy*. 102: 103–26.
- Werden, G.J. 1987. "The Law and Economics of the Essential Facility Doctrine." *St. Louis University Law Journal*. 32: 433–80.
- Werden, G.J. 2006. "Identifying Exclusionary Conduct under Section 2: The "No Economic Sense" test." *Antitrust Law Journal*. 73: 413–34.
- Whinston, M.D. 1990. "Tying, Foreclosure and Exclusion." *American Economic Review*. 80: 837–59.
- Whinston, M.D. 2006. *Lectures on Antitrust Economics*. Cambridge, MA: MIT Press.
- Whish, R. and D. Bailey. 2015. *Competition Law*. Oxford: Oxford University Press.
- White, C.L. and H. Good. 2009. "All Together Now? Evolving Antitrust Approaches to Bundled Discounting." *Journal of Health & Life Sciences Law*. 2: 47–71.
- Wright, J.D. 2009. "Exclusive Dealing and Entry, When Buyers Compete: Comment." *American Economic Review*. 99: 1070–81.

- Wright, J.D. 2013. "Simple but Wrong or Complex but More Accurate? The Case for an Exclusive Dealing-Based Approach to Evaluating Loyalty Discounts." Speech delivered at the Bates White 10th Annual Antitrust Conference, Washington, DC, on 3 June 2013.
- Wright, J.D. 2015. "The Supreme Court Should Grant Certiorari in *FTC v. McWane*." *George Mason University Law and Economics Research Paper Series No. 15-57*.
- Yamey, B.S. 1972. "Predatory Price Cutting: Notes and Comments." *The Journal of Law and Economics*. 15: 129-42.
- Yong, J.-S. 1996. "Excluding Capacity-Constrained Entrants Through Exclusive Dealing: Theory and an Application to Ocean Shipping." *The Journal of Industrial Economics*. 44: 115-29.
- Yoshida, Y. 2000. "Third-Degree Price Discrimination in Input Markets: Output and Welfare." *American Economic Review*. 90: 240-46.