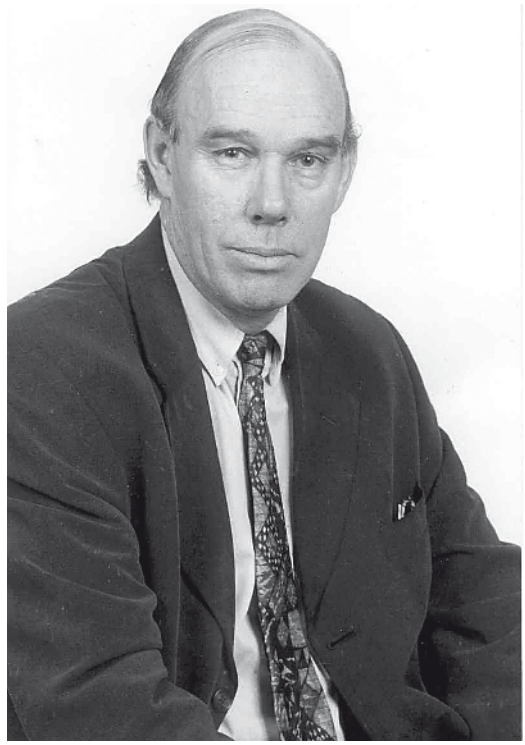
Growth and Economic Development



A.P. Thirlwall

Growth and Economic Development

Essays in Honour of A.P. Thirlwall

Edited by

Philip Arestis

University Director of Research, Cambridge Centre for Economic and Public Policy, Department of Land Economy, University of Cambridge, UK

John McCombie

Director, Cambridge Centre for Economic and Public Policy, Department of Land Economy, University of Cambridge and Fellow in Economics, Downing College, Cambridge, UK

and

Roger Vickerman

Jean Monnet Professor of European Economics, University of Kent, Canterbury, UK

Edward Elgar
Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
Glensanda House
Montpellier Parade
Cheltenham
Glos GL50 1UA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Cataloguing in Publication Data
Growth and economic development : essays in honour of A.P. Thirlwall /
edited by Philip Arestis, John McCombie, and Roger Vickerman.
p. cm.
Includes bibliographical references and index.
1. Economic development. 2. Development economics. I. Arestis, Philip, 1941– II. McCombie, J. S. L. III. Vickerman, R. W. (Roger William) IV. Thirlwall, A. P.

HD82.G753 2006
338.9—dc22

2006010516

ISBN 978 1 84376 878 4

Typeset by Manton Typesetters, Louth, Lincolnshire, UK
Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

Contents

About the contributors

Philip Arestis is University Director of Research, Cambridge Centre for Economic and Public Policy, Department of Land Economy, University of Cambridge, UK, Adjunct Professor of Economics at the University of Utah, USA and Senior Scholar at the Levy Economics Institute, New York, USA. He is also Visiting Professor, University of Leeds, UK, and School of Oriental and African Studies (SOAS), University of London, UK. He was a member of the *Economics and Econometrics* RAE panels in 1996 and 2001. His research interests are in the areas of Macroeconomics, Monetary Economics, Applied Econometrics, Political Economy and Applied Political Economy. He has published as sole author or editor, as well as co-author and co-editor, a number of books, and in academic journals.

John Cornwall is Professor Emeritus at Dalhousie University, Halifax, Canada and Associate Member, Cambridge Centre for Economic and Public Policy, University of Cambridge, UK. He has authored and edited several books on theoretical and applied macroeconomics including *Growth and Stability in a Mature Economy* (1972), *Modern Capitalism* (1977), *Economic Recovery for Canada* (with Wendy Maclean; 1984), *The Theory of Economic Breakdown* (1990) and *Capitalist Development in the Twentieth Century* (with Wendy Cornwall; 2001). His articles have appeared in such journals as *Review of Economics and Statistics, Quarterly Journal of Economics. American Economic Review, Kyklos, Economic Journal, Journal of Post-Keynesian Economics,* and *Journal of Economic Issues.*

Paul Davidson is Editor of the *Journal of Post Keynesian Economics*. He is currently a Visiting Scholar at the Bernard Schwartz Center for Economic Policy Analysis of the New School in New York City. Previously he occupied the Holly Chair of Excellence of Political Economy at the University of Tennessee. He has also had professorships at Rutgers University and the University of Pennsylvania. He has also had visiting appointments at Bristol University and Cambridge University in the UK and the Institute for Advanced Studies in Vienna. He was also the Assistant Director of the Economics Division of the Continental Oil Company.

Valpy FitzGerald is Professor of International Economics and Finance, Director of the Finance and Trade Policy Centre and Professorial Fellow of St Antony's College at Oxford University. His recent books include *Global Markets and the Developing Economy* (Basingstoke: Palgrave, 2003) and *The Transmission of Economic Ideas in Latin America* (Basingstoke: Palgrave, 2005). He is currently engaged on research into long-run economic development in Latin America, the determinants of portfolio investment flows towards emerging markets and the relationship between taxation and income distribution.

Heather D. Gibson is Head of the Special Studies Division in the Economic Research Department of the Bank of Greece. Among her duties is responsibility for the Bank's *Economic Bulletin* as well as the publication of their working paper series. Before joining the Bank of Greece, she was a Lecturer in Economics at the University of Kent. Her main interests include European integration, comparative financial systems, international financial flows and exchange rate systems. She has authored several books and has published widely in journals.

G.C. Harcourt was born in Melbourne, Australia. He is a graduate of Melbourne and Cambridge Universities. He is Emeritus Reader in the History of Economic Theory, Cambridge University (1998), Emeritus Fellow, Jesus College, Cambridge (1998) and Professor Emeritus, Adelaide University (1988). Harcourt is author/editor of 24 books, and over 230 articles or chapters in learned journals and edited volumes.

Mohammed Nureldin Hussain was Division Chief of the African Development Bank's Research Division from 2002 until his tragic death in July 2005. He held a PhD in Economics from the University of Kent at Canterbury. He acquired 25 years of experience in development economics through teaching, research and application. His work experience includes 9 years as a development practitioner with the African Development Bank, and over 16 years as teaching assistant, lecturer, and associate professor of economics in several universities. His writings have been published widely, covering issues of relevance to Africa's development with original and recognised contributions to development theory and practice.

John McCombie is Director of the *Centre for Economic and Public Policy*, University of Cambridge and Fellow in Economics at Downing College, Cambridge. He has previously held appointments in the Departments of Economics at the University of Hull and the University of Melbourne. His major research interests are in regional economics, Keynesian macroeconomics and the causes of growth disparities (which includes balance-of-payments constrained growth

models and a critique of the measurement of technical change using the neoclassical aggregate production function). He has published and co-edited several books, as well as published widely in economic journals.

Miguel A. León-Ledesma (PhD, La Laguna) is a Senior Lecturer in Economics at the University of Kent. His main research areas are Macroeconomics, Applied Econometrics, and Economic Growth. Together with Tony Thirlwall he developed the concept of the 'endogenous natural rate of growth'. His research has been published in several journals such as *Cambridge Journal of Economics, Journal of Macroeconomics, Journal of International Money and Finance* and *Scottish Journal of Political Economy*. He has been consultant for the Central Bank of the Dominican Republic and the European Central Bank.

Mark Roberts is a Fellow in Economics at New Hall and an Affiliated Lecturer in Regional Economics at the University of Cambridge. He is also a member of the *Cambridge Centre for Economic and Public Policy*, where he is currently engaged in collaborative work on the determinants of spatial disparities in rates of productivity, efficiency and technical change within the EU. His other research interests include the theory and empirics of economic growth from a Kaldorian perspective. He has published in several journals including *International Review of Applied Economics*, *Journal of Regional Science*, *Regional Studies* and *The Manchester School*, as well as in a range of edited books.

Malcolm Sawyer is Professor of Economics, University of Leeds. He is managing editor of *International Review of Applied Economics*, managing co-editor of *International Papers in Political Economy* and joint editor of *Journal of Income Distribution*. He is the editor of the series *New Directions in Modern Economics* published by Edward Elgar. He is the author of 11 books, has edited 18 books and has published nearly 150 articles and chapters.

Mark Setterfield is Professor of Economics in the Department of Economics at Trinity College, Hartford, Connecticut and Associate Member, Cambridge Centre for Economic and Public Policy, University of Cambridge, UK. His main research interests are macrodynamics (with a particular focus on concepts of path dependence) and Post-Keynesian economics. He has written and edited several books, and has published in numerous journals including the *Cambridge Journal of Economics*, *Journal of Post Keynesian Economics*, *European Economic Review*, *Review of Political Economy*, *Journal of Economic Issues* and *The Manchester School*.

Ajit Singh is Professor of Economics at Cambridge University and Senior Fellow at Queens' College. His research spans three areas: (i) The theory of the

firm and industrial organisation including the role of the stock market and the relationship between law, finance and industrial development; (ii) De-industrialisation and structural change in developed and developing countries; and (iii) Economic policy and development. He has written several books and published extensively in leading economic journals. He has been a senior economic advisor to the governments of Mexico and Tanzania and a consultant to various UN developmental organisations.

Roberto Tamborini is Head of the Economics Department at the University of Trento, Italy and was appointed in 2001. He did his first degree (Graduation in Economics) at the University of Modena in 1982 and subsequently did postgraduate work at the University of Cambridge (M.Phil. in Economics) and at the European Institute, Florence (PhD in Economics). He has undertaken a period of research at the University of California at Berkeley, USA and has been a Visiting Scholar at Boston University, USA and IDEFI-CNRS, Sophia-Antipolis, Nice, France. His research in recent years has mainly concentrated on macroeconomic theories and policies, money and financial markets, and policies and institutions.

Ferdinando Targetti has a *Laurea* in Economics and Business Studies from Bocconi University in Milan and has done postgraduate work at Cambridge University. Since 1984 he has been Professor of Economic Policy at the University of Trento, and since 2001 he has been the Director of the postgraduate School of International Studies. His research interests include Italian economic development, the theory of value and distribution, the history of economic thought (Kaldor and Kalecki), international economics, and Italian fiscal reform and budget policy. He has published widely in journals, books and edited volumes. Between 1996 and 2001 he was a member of the Italian Parliament and the Public Finance Commission.

John Toye was educated at Cambridge University, Harvard University and SOAS at London University. He has been professor of development economics at (successively) the universities of Wales, Sussex and Oxford, where he also directed Centres or Institutes of development studies. He has worked for short periods in the British civil service (HM Treasury), the private sector (as director of CRU Ltd) and the United Nations (as a Director of UNCTAD). He has also been consultant for DFID (formerly ODA), the World Bank, the European Commission, ILO, WHO, the Commonwealth Secretariat and the US Overseas Development Council. He is the author of seven books and numerous journal articles.

Euclid Tsakalotos is a Professor in the Department of International and European Economic Studies at Athens University of Economics and Business.

Previously he was a Lecturer in Economics at the University of Kent. His interests lie in a number of areas of economics, including macroeconomics and European integration, financial and industrial economics, labour economics and political economy. The common theme in much of his work is the importance of institutions and their impact on economic performance. He has published articles, books, and contributions to edited volumes, with his papers having appeared in a number of journals.

Roger Vickerman is Professor of European Economics and Head of the Department of Economics at the University of Kent, Canterbury, UK. He studied Economics at Clare College, Cambridge, has a DPhil in Economics from the University of Sussex and an honorary doctorate in Economics from the Phillipps-Universität Marburg. His main research interests are the relationship between transport (especially infrastructure), regional development and integration in the European Union; and the role of migration and labour mobility in integration. He was a member of the Standing Advisory Committee on Trunk Road Assessment (SACTRA) in the UK contributing to its major report on *Transport and the Economy*, and has recently been involved in a number of studies on transport developments in the EU. He has served as an advisor to committees of both Houses of Parliament and acted as a consultant to the European Commission, various UK government departments and regional and local government authorities.

Introduction

Philip Arestis and John McCombie

Where does one begin to review Tony Thirlwall's substantial contribution to economics? We considered starting chronologically with his early pioneering work in regional economics and labour economics. There he examined, *inter alia,* regional differences in unemployment and how disequilibrium in the labour market affects inflation. However, Tony has been so productive that short of an intellectual biography, it is difficult to do justice to him. Consequently, we are not going to attempt a comprehensive review, but rather let the contributions to this *Festschrift* bear eloquent testimony to his influence on the profession. Nevertheless, we felt it appropriate to write a few words by way of introduction to the volume, albeit it as a brief Cook's tour of Tony's contributions. It also follows that we have been necessarily somewhat eclectic. Fortunately, there are now two volumes of Tony's collected essays that provide a useful and convenient reference to his work, although, having been published in Thirlwall 1995 and 1997, it is clearly time for a third volume! Even a casual perusal of these volumes will quickly establish the wide variety and scope of his output.

Tony Thirlwall is best described as a Keynesian applied economist, a category he would probably not object to. Indeed, he considers the term 'unreconstructed Keynesian' which he has used to describe himself, not to be a pejorative title, but more an accolade! He is an applied economist, but not because he eschews theory; far from it. As we document below, he has made major contributions to economic theory. He is applied because he believes that to be of any importance: theory must be related to, in that hackneyed but nevertheless relevant phrase, the 'real world'. He has no time for esoteric theory, no matter how elegant. As Tony says, 'In my own research in, and writing in later years, I have always tried to treat economics as a moral science – as a branch of ethics – in the Keynesian (Cambridge) tradition, choosing policy issues of public concern and using economic theory primarily as an aid to policy analysis' (Thirlwall, 1995, p. ix).

He has a consistent view of how the macro economy works, which takes its inspiration from Keynes and comes from a carefully reading of the *General Theory* – how many economists today can say that? As Tony puts it, 'If Milton Friedman and his disciples had ever properly understood the *General Theory,*

how could they claim that "money doesn't matter in Keynes"? If Robert Lucas and his followers knew their Keynes, how could they possibly maintain that Keynes's macroeconomic model cannot explain stagflation?' (1995, p. x).

One of the pleasures of reading any of Tony's papers is that he writes clearly as well as persuasively. This is true of both his technical and his more general papers. As far as the latter is concerned it is difficult to do better than to mention his 1981 paper *Keynesian Employment Theory is not Defunct* together with his 1993 paper, *The Renaissance of Keynesian Economics*. The first, as its title suggests, is a clear statement of the insights Keynes still has for the functioning of the economy, while the latter article shows how subsequent economic events have amply justified Tony's earlier arguments. Both these papers should be compulsory reading for first year economics students, if only to show them that there is more to macroeconomics than may be gleaned from the standard textbooks such as Mankiw (2002).

It was not only through his writings that Tony kept the Keynesian, or rather Keynes's, message alive during the years of the onslaught from monetarism and new classical economics. He organised eleven Keynes seminars at the University of Kent between 1972 and 1993 that touched on a wide variety of Keynesian topics. The seminars were packed out, and a wider audience was reached through the publication of the various papers. In the 1970s, the notable contemporaries of Keynes (Roy Harrod, Richard Khan, Joan Robinson, and Nicholas Kaldor) were present at one time or another at the seminars. This was perhaps the only time many of the younger generation of Keynesian economists had a chance to meet these giants of the profession and to see them in action.

Apart from his work on short-run macroeconomics (multiplier analysis, unemployment, inflation etc.) much of Tony's research is focussed on the twin themes of growth and development and the open economy. One can see in some of his work the influence of the late Professor Lord Kaldor, and much of Tony's work is in what may be best described as the Kaldorian tradition (see, for example, Kaldor, 1995). Tony is Kaldor's literary executive and wrote a definitive biography of him, which not only clearly explains Kaldor's economic *Weltanschauung*, but is a good read into the bargain (Thirlwall, 1987). With Fernando Targetti, Tony also compiled a collection of Kaldor's most influential papers published in 1989 as the *Essential Kaldor* and they also edited (together with C. Filippini) Kaldor's Mattioli lectures.

Kaldor (and, before him, Gunnar Myrdal) had long argued that it was not possible to explain why some countries are rich and others poor without understanding that growth occurred by a process of cumulative causation. Feedback effects mean that there are powerful forces that perpetuate the fast growth rate of the successful countries and cause slow growing countries to be caught in a vicious circle of low productivity and low output growth. The early Solow growth model, with its emphasis on convergence and its artificial (and not very

illuminating) dichotomy of growth into that caused by the growth of factor inputs and that resulting from technical change, was not seen as a productive approach. There is no point in distinguishing between shifts of the production function and movements along the production function and as the two movements are inextricably related. (If the overwhelming economic forces are for convergence, how does one explain the fact that two centuries ago the ratio in per capita income between the richest and the poorest areas were about 1:3 and now they are 1:70?)

At the heart of Kaldor's explanation, long before the development of endogenous growth theory, was that much of technical progress was induced by the growth of output. This was formulated in his famous technical progress function, of which the Verdoorn law can best be regarded as an empirical counterpart. The Verdoorn law is the statistical relationship found between productivity and output growth. A statistically significant (Verdoorn) regression coefficient is interpreted as evidence of increasing returns to scale (of both the dynamic and static variety) and induced technical change. Tony found much to commend in Kaldor's emphasis on the importance of beginning with empirical relationships and 'Kaldor's three growth laws', of which the Verdoorn law is one.[1] He organised a symposium whose proceedings were published in the 1983 edition of *Journal of Post Keynesian Economics* and did much to stimulate interest in these ideas. Nearly all of Kaldor's later key writings were of a verbal nature, and were publications of addresses or lectures he had given (see, for example, his 1979 'Case for Regional Policies' or his 1982 'Irrelevance of Equilibrium Economics'). It was Tony who, along with a then PhD student of his, Robert Dixon, formalised in 1975 Kaldor's growth analysis in what has become the canonical cumulative causation model. Others have later extended the model by, *inter alia*, including the supply side, but even after 30 years the original article remains essential reading for any student of the growth process. He also formalised Kaldor's two-sector model, neatly capturing the essence of Kaldor's insights on the importance of the growth of sectoral demand in the economic growth process, and also clarifying the argument (Thirlwall, 1986).

Tony has long been interested in the open economy, and indeed he wrote a major textbook on the subject that went to four editions, the last co-authored with Heather Gibson in 1993. His interests in growth and the open economy came together in his famous balance-of-payments constrained growth model and what came to be known as 'Thirlwall's law'. Mohammed Hussain's contribution to this volume tells how Tony first introduced the law to his postgraduate students at the time when he was still drafting the paper. It is now far too late to call it 'Thirlwall's rule', as Mohammed suggests was Tony's modest preference. But it is doubtful if Tony had any idea of just what impact the model, published in 1979 in the *Banca Nazionale del Lavaro Quarterly Review,* would have on the profession. There has been a proliferation of papers since 1979,

testing the law and elaborating its theoretical basis (see the collection of papers in McCombie and Thirlwall, 2004). It is fair to say that it now stands as the major alternative explanation to the neoclassical model of why growth rates differ.

The simplest version of the law states that the rate of growth of output consistent with the balance of payments being in equilibrium is given by $y_B = x/\pi = \varepsilon z/\pi$, where x is the growth of exports, π is the income elasticity of demand for imports, ε is the income elasticity of demand of world demand for exports and z is the growth of world income. The explanation of this law is relatively straightforward. If changes in relative prices have little effect on the growth of exports and imports, and the evidence suggests that this is the case, then any growth of income faster than y_B will have to be financed from abroad by a growing inflow of capital. The overseas debt-to-income ratio will grow, but this will be unsustainable in the long run. The only way to prevent this occurring is to reduce the growth of imports, and the only effective means to accomplish this is to reduce the rate of growth of output. If, as a result, this is below the growth of productive potential, then the growth rate is 'balance-of-payments constrained'. The result is a lower rate of capital accumulation and induced technical change than would otherwise be the case, with increasing overt and/or disguised unemployment. In an important paper, Tony, with Miguel León-Ledesma, (2002) has shown that the natural rate of growth is indeed endogenous and not immutable.

Differences in growth rates, in this schema, are predominantly caused by international differences in the values of ε and π, which reflect all aspects of non-price competitiveness – factors which change relatively slowly over time. Of course, the model and the argument can be qualified in a number of ways, and future lines of research that suggest themselves, for example, elaborating and analysing the determinants of non-price competitiveness; what affects it and how it changes over time. But the essential insight of the simple model still remains and it is remarkable how well the model stands up to statistical testing (McCombie and Thirlwall, 2004). The model amply fulfils one of Tony's requirements that a good model 'explains a lot from a little'. Its policy implications are far reaching. A major problem for a developing country is to acquire sufficient foreign exchange earnings to pay for especially the capital goods required for development etc. and at the same avoid the balance-of-payments constraint. But if a large number of countries target the same export sector or product, the world income elasticity of demand of an individual country's exports in that sector may be low, even though world demand for the product overall is growing rapidly.

In 2000, Tony gave a series of lectures at the National University of Mexico on the nature of economic growth. In the resulting short book, *The Nature of Economic Growth*, he clearly and in a successfully non-technical way lays out

his views of, in the words of the subtitle, *An Alternative Framework for Understanding the Performance of Nations*. This is, again, another publication that should be on every undergraduate's reading list.

The tile of this Festschrift, *Growth and Development*, was chosen with care as the other great interest of Tony is in the whole field of economic development. His monumental book, *Growth and Development, with Special Reference to Developing Countries* has established itself as the leading textbook in the area and is now in its eighth edition, the first being published in 1972. It bears witness to Tony's encyclopaedic knowledge of development issues. (As Roger Vickerman documents in the next article, Tony has built up a large and successful MPhil in Development Economics at the University of Kent with many of his students going on to important posts in development agencies etc.) Much of his research, such as the balance-of-payments constrained growth, has immediate implications for development as Mohammed Hussain shows in Chapter 1. Tony has also been adviser to many developing countries, with strong links to the African Development Bank. He has, at times, been controversial, daring to criticise IMF policies and to raise the heretical view that possibly, just possibly, unilateral free trade and unfettered capital liberalisation, as well as exchange rate unification in all circumstances, may not be in the best interests of a developing country.

Tony's desire to further understanding of the growth process shows no signs of abating. He has recently examined the role of trade liberalisation on economic growth (Thirlwall and Santos-Paulino, 2004) and will no doubt produce many more important insights in the years to come. We noted at the beginning of this Introduction that we would prefer to let the papers contributed to this volume bear witness to his significant contribution to the economics; so let us turn to these.

Mohammed Nureldin Hussain was a pupil, collaborator, and friend of Tony's and it is with sadness that we record his untimely death after he had submitted his chapter, *The Implications of Thirlwall's Law for Africa's Development Challenges*, for this *Festschrift.* Mohammed Hussain provides a fascinating pen-portrait of Tony, as well as an intriguing insight into the development of Thirlwall's law, prior to the publication of the famous *Banca Nazionale del Lavoro Review* article in 1979. He shows the relevance of the law and its extensions to the contemporary debate on Africa's development issues. Thirlwall's law proves to be highly relevant for the setting of African development priorities related to: (i) poverty reduction; (ii) growth and debt sustainability; (iii) globalisation and the income dimension of global competition; and (iv) the development effectiveness of foreign aid in relation to the issue of transformation. Hussain also demonstrates that Thirlwall's Law has a multitude of analytical capabilities and he concludes with some policy implications related to a long-term strategy of socio-economic development in Africa.

Mark Setterfield in *Thirlwall's Law and Palley's Pitfalls: A Reconsideration* answers a challenge to the balance-of-payments equilibrium growth model posed by Palley (2002). Palley considers that in Thirlwall's schema, with the growth of demand being determined by the growth of exports via Thirlwall's Law, and the growth of the supply of output being determined by the growth of productivity and the growth of the labour force (through the Verdoorn law), the model is over-determined. This is true *if* the growth of supply forthcoming is interpreted as the growth of productive potential or the natural rate of growth, as Palley maintains.[2] One solution Palley proposes is to have the income elasticity of imports determined by the supply conditions. Thus, if the balance-of-payments growth rate is slower than the growth of potential supply, for some reason the income elasticity of demand for imports falls, bringing the two growth rates into equality. In this scenario, there cannot be, by definition, any balance-of-payments constraint. Mark Setterfield, however, proposes a more plausible scenario. With a lower growth rate and a depressed labour market, the size of the Verdoorn coefficient falls. As the Verdoorn coefficient captures both increasing returns to scale and induced technical change, this causes the growth of productivity to fall. Thus, while the labour force may be fully employed, the growth of productivity (and hence real wages) is below the level that would occur if export growth were faster. This is very much in accord with Tony Thirlwall's contention that the natural rate of growth is endogenous.

Geoffrey Harcourt, on the other hand, concentrates on a somewhat narrower aspect of Thirlwall's work in his chapter *On Specifying the Demand for Imports in Macroeconomic Models*. Although specifying imports as a function of gross expenditure (i.e., expenditure including that on imports) rather than national income had been briefly mentioned in the past, it was Tony Thirlwall with Charles Kennedy who first fully articulated the case for using expenditure. Geoff Harcourt, however, suggests that the conventional use of income, rather than expenditure might have been right all along, but for the wrong reason and thus presents a challenge for Tony Thirlwall ('I shall be the first to cheer if Tony shows that I am now wrong').

Paul Davidson in *Post Keynesian Analysis and the Open Economies of the Twenty-first Century* shows that while the *General Theory* was essentially about a closed economy, it nevertheless contains important insights for the global economy, a view with which Tony certainly agrees. Paul Davidson cites chapter and verse, arguing persuasively that many of the policies advocated in the *General Theory* are equally relevant to today's more open economies. When there are unemployed or underemployed resources, an increase in exports by one country will merely lead to a fall in demand and employment in the importing economy. This can only be avoided by a co-ordinated expansion of demand as flexible exchange rates are inadequate to ensure full employment in all countries. He shows that Thirlwall's law, which reflects the dynamic Harrod foreign trade

multiplier, is a simple yet powerful explanation for the understanding of growth rate differences in both advanced and less developed countries. 'Consequently', Paul Davidson concludes, 'Thirlwall's law demonstrates that international financial payment imbalances can have severe real consequences, i.e. money is never neutral in an open economy'.

The next three chapters move away from the open-economy model and consider growth in a closed economy framework. The relationship between short-run cycles and long-run growth is a central theme in Thirlwall's contributions to growth economics, who has shown that the natural rate of growth is endogenous. In the next chapter entitled *Cycles, Aggregate Demand and Growth,* Miguel León-Ledesma discusses the possible relations between cycles, aggregate demand shocks, and growth. He further proposes a simple time-series method for analysing this statistical relation that endogenizes the impact of cycles on the definition of trend output. He then applies this method to analyse the cases of the US, Germany and the UK and finds that cycles do seem to have a strong impact on trend output, but this impact is different for the three economies in question.

Mark Roberts revisits a debate concerning the compatibility of Kaldor's writings on the cumulative causation nature of economic growth and his well-known critique of equilibrium economics. Mark Setterfield (1997) has pointed out that the formalisation of the cumulative causation model along the lines of Dixon and Thirlwall is not truly path-dependent. This is because the solution of the model depends only upon the exogenous parameters. Setterfield constructed an extended model to overcome this shortcoming, although, as he later admits, this was not entirely successful. While it has been argued by some that it is not possible to formally (mathematically) model Kaldor's vision of the growth process, Mark Roberts disputes this. He builds on Setterfield's approach to construct a model that is, as Roberts puts it, 'consistent with the *spirit* of Kaldor's historical growth process'. This involves the adoption of an 'open systems-*ceteris paribus*' approach to modelling.

The chapter by Philip Arestis and Malcolm Sawyer, *Endogenous Growth Theory: A Partial Critique* begins by pointing to the many contributions of Tony Thirlwall. His writings on growth theory with his emphasis on the role of demand, balance-of-payments constraints, endogeneity of the natural rate and the role of sectoral analysis have made major contributions to understanding the growth process and also provide a continuing (sometimes implicit, sometimes explicit) critique of endogenous growth theory. It is, thus, very fitting that Philip Arestis and Malcolm Sawyer indulge in a critique of endogenous growth theory. The critique is partial in the sense of not being comprehensive which is precluded by space considerations. (But no doubt Tony Thirlwall himself would be partial to it!) They review briefly the endogenous growth theory, and concentrate on the returns to scale assumption and its implications. This is followed

by a discussion on the approach to investment and productivity growth in endogenous growth theory. A brief review of the relevant empirical evidence is offered before they conclude that the critique they offer sits very comfortably with Tony Thirlwall's numerous contributions on this matter.

John McCombie in *The Nature of Economic Growth and the Neoclassical Approach: More Questions than Answers?* also critically reviews the neoclassical approach to economic growth. For reasons of space, this chapter is eclectic and deliberately does not concern itself with the Keynesian or Kaldorian approach. He considers the role of capital accumulation and technical progress in the Harrod–Domar model and the Solow model and concludes, like Richard Nelson and others, that these approaches do not tell us much that we did not already know. The new growth theory is shown to incorporate certain theory-dependent assumptions that are simply there to give the model a steady-state solution. Finally, McCombie draws upon work he has done with his colleague, Jesus Felipe, that questions the whole notion of the aggregate production function (a concept which is the *sine qua non* of neoclassical macroeconomics) and whether it can be tested. He shows that, in this light, the Mankiw–Romer–Weil (1992) growth model has little or no explanatory value.

John Cornwall is more concerned with important issues of the short run, rather than economic growth *per se.* His chapter, *On the Core of Macroeconomics*, can be regarded as falling into two parts. The first part provides a critique of mainstream macroeconomics theory. It argues that the New Keynesian Macroeconomics suffers from serious errors of commission and omission, ultimately traceable to its failure to shed the neoclassical roots of both its short-run and long-run analysis. These errors lead to serious inconsistencies in its theoretical structure and its inability to explain the broad historical tendencies of macroeconomic development. In the second part of the chapter, Cornwall develops a macroeconomic theory based on macroeconomic principles associated with Keynes' *General Theory,* an approach with which Tony Thirlwall has a great deal of sympathy. The result is an extended Keynesian core theory capable of describing in greater detail the short-run and long-run macroeconomic processes at work in advanced capitalist economies, and of providing an account of the historical development of modern economies.

One of the major macroeconomic concerns in Europe at the moment is the functioning (or lack of functioning) of the Stability and Growth Pact. While sanctions were brought to bear on a small country, Portugal, that breached the Pact regulations, once this was perpetrated by two of the major players, France and Germany, the Stability and Growth Pact was suspended. Roberto Tamborini and Fernando Targetti outline the problems inherent in the Pact and concede that 'it was obtained with significant damage to the already fragile structure of the Community institutions'. The solution, they contend, is not to be found in further tinkering with the Pact. What is required is a more radical approach: the

creation of a European Confederation Treasury, which would have the substantial fiscal responsibilities that many fully integrated federal systems have. They admit hat this is not politically feasible, at least at the present. Consequently, they outline a 'second best' solution of reforming the budgetary rules and setting up a Budgetary Authority without changing the 'distribution of powers'.

Keynesian economists have long sought to promote a framework for economic decision-making that allows for the clash of ideas to have some impact on outcomes. They have not seen democratic accountability and discretion as a problem to be overcome through various devices such as independent central banks, budgetary rules and so on, but rather as part of the answer to good economic policy-making. In their chapter Heather Gibson and Euclid Tsakalotos argue that this approach is absent from the institutional framework and the policies of the EU, which are far from neutral with respect to the choices member states make between the two poles of a liberal capitalist economy and a more institutional economy. The pressure is to keep as close as possible to the liberal pole. They support this argument by examining both the macroeconomic framework and the financial framework and discuss the implications for the new EU member states.

The remaining chapters reflect Tony Thirlwall's long abiding interest in development issues and problems. In *Competition and Competition Policy in Emerging Markets: International Development Decisions,* Ajit Singh examines the role of competition policy in emerging markets from a developmental and international perspective. It comes to a number of important conclusions. Contrary to conventional wisdom, evidence suggests that the intensity of competition in leading emerging markets is certainly no less than that observed in advanced countries. He shows that analysis and evidence indicates that maximum competition is not necessarily optimal in terms of dynamic efficiency. There is little evidence to indicate that the current international merger wave will enhance global economic efficiency. An implication is that developing countries need a competition policy today, because of (i) privatisation and deregulation, and (ii) the huge international merger movement. However, the current competition policies in the US and the European Union are unsuitable for developing countries. Countries at different levels of development and governance capacities require different types of competition policies. The chapter concludes with a proposal for a development-oriented international competition authority to control anti-competitive conduct and growth by mergers of large multi-nationals.

Next, Valpy FitzGerald takes up Thirlwall's two models of saving behaviour in developing countries in *Models of Saving, Income and the Macroeconomics of Developing Countries in the Post-Keynesian Tradition*, and expresses them formally in terms of the modern inter-temporal and open-economy frameworks. Reinterpreting Thirlwall's Keynesian view on the relationship between per

capita income and the aggregate savings rate as a wealth effect in the intertemporal framework justifies his qualitative insights into financial liberalisation. Placing Thirlwall's Kaldorian model of the link between wages, inflation and saving within the modern open-economy framework clarifies the role of the exchange rate. Finally the chapter extends these ideas by including an autonomous investment function to reveal the medium term effects of the real exchange rate on real wages and thus investment and employment in a post-Keynesian approach to the adjustment process.

John Toye, as the title of his chapter *The Influence of Keynes on Development Economics* suggests, returns to John Maynard Keynes, who exercised considerable influence on the formative years of development economics through his disciples. Key issues were the relevance of disguised unemployment to developing countries and the supply side policies that would have to complement Keynesian demand management. Although Michael Kalecki made a greater contribution to structuralist economics, Keynesian accounting formats still dominate macroeconomic development models. The legend of Keynes – his policy advocacy and intellectual innovation – also inspired emulation by a younger generation of development economists, both in 'revolutionising' economic theory and in promoting the Keynesian version of the international economic architecture. Ironically, Toye suggests, his writings on Soviet development were neglected, and Keynes's modern critics unknowingly used his own arguments to try to discredit his influence.

We are extremely grateful to Edward Elgar for his encouragement to put together this volume in honour of A.P. Thirlwall, and to his efficient staff for making this project possible. We are of course equally grateful to the authors for their contributions to this volume. The University of Kent Economics Department very generously hosted a conference to celebrate A.P. Thirlwall's numerous contributions not merely to this department but also to the discipline of economics. We thank them all for their input to this book and relevant celebrations.

NOTES

1. The others showed the importance of the manufacturing sector as the engine of growth and the importance of the intersectoral transfer of labour in explaining productivity growth disparities.
2. This misunderstands the nature of a binding constraint. The balance-of-payments constrained equilibrium growth rate, as it is by definition a constraint on economic growth, is below the growth of productive potential (the natural rate of growth). The latter is the *maximum* growth of output that could be achieved, given supply side factors, with the necessary accompanying growth of exports. In other words, it is the maximum possible growth of output that would occur if there has not been any demand constraints in the past. See Thirlwall (2001). The result is that the balance-of-payments constrained equilibrium growth rate leads to increasing unemployment

and/or disguised unemployment (especially in the service and public sector), increased emigration and/or reduced immigration and a slower rate of capital accumulation. There will also be a fall in the rate of growth of productivity through increasing x-inefficiency, as labour markets are not tight, and also through the Verdoorn effect. As Mark Setterfield points out, the size of the Verdoorn coefficient itself may fall.

However, the growth of productive potential is a somewhat ambiguous concept. The rate of growth at which the supply constraint becomes binding will depend upon the past performance of the economy. In other words, there is path-dependent process in operation. (See the chapter by Mark Roberts in this volume.) Nevertheless, it is unlikely that the growth of productive potential will necessarily converge to the balance-of-payments equilibrium rate, in the sense that if a sustained increase in exports occurs, this will eventually lead to an increase in the long-run growth of output. All this is either implicit, or explicit, in the literature on the balance-of-payments constrained growth model (see, for example, Thirlwall, 1979, 2001, and McCombie and Thirlwall, 2004). The irony is that one of the putative solutions that Palley himself proposes to overcome what he sees as an 'inconsistency' in the model is along the lines mentioned above. Mark Setterfield gives an insightful discussion of these issues.

REFERENCES

Dixon, R.J. and Thirlwall, A.P. (1975), 'A Model of Regional Growth-rate Differences Along Kaldorian Lines', *Oxford Economic Papers*, **27**, 201–14 (reprinted in Thirlwall, 1997).

Kaldor, N. (1979), 'The Case for Regional Policies', *Scottish Journal of Political Economy*, **17**, 337–48.

Kaldor, N. (1982), 'The Irrelevance of Equilibrium Economics', *Economic Journal*, **82**, 1237–55.

Kaldor, N. (1995), *Causes of Growth and Stagnation in the World Economy*, Raffaele Mattioli Lectures, Filippini, C., Targetti, F. and Thirlwall, A.P. (eds), Cambridge: Cambridge University Press.

León-Ledesma, M.A. and Thirlwall, A.P. (2002), 'The Endogeneity of the Natural Rate of Growth', *Cambridge Journal of Economics*, **26**, 441–59.

McCombie, J.S.L. and Thirlwall, A.P. (2004), *Essays on Balance of Payments Constrained Growth: Theory and Evidence*, London: Routledge.

Mankiw, N.G. (2002), *Macroeconomics*, New York: Worth Publishers.

Mankiw, N.G., Romer, D., and Weil, D.N. (1992), 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, **107**, 407–37.

Palley, T.I. (2002), 'Pitfalls in the Theory of Growth: An Application to the Balance-of-Payments Constrained Growth Model,' in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Setterfield, M. (1997), *Rapid Growth and Relative Decline: Modelling Macroeconomic Dynamics with Hysteresis*, London: Macmillan.

Targetti, F. and Thirlwall, A.P. (1989), *The Essential Kaldor*, London: Duckworth.

Thirlwall, A.P. (1979), 'The Balance of Payments Constraint as an Explanation of International Growth Rate Differences', *Banca Nazionale del Lavoro Quarterly Review*, **128**(791), 450–53 (reprinted in Thirlwall, 1997, and McCombie and Thirlwall, 2004).

Thirlwall, A.P. (1981), 'Keynesian Employment Theory is not Defunct', *Three Banks Review*, September (reprinted in Thirlwall, 1997).

Thirlwall, A.P. (ed.) (1983), 'Symposium on Kaldor's Growth Laws', *Journal of Post Keynesian Economics*, **5**.
Thirlwall, A.P. (1986), 'A General Model of Economic Growth and Development Along Kaldorian Lines', *Oxford Economic Papers*, **38**, 199–219 (reprinted in Thirlwall, 1995).
Thirlwall, A.P. (1987), *Nicholas Kaldor*, Brighton: Harvest Wheatsheaf.
Thirlwall, A.P. (1993), 'The Renaissance of Keynesian Economics', *Banca Nazionale Del Lavoro Quarterly Review*, **0**(186), 327-37 (reprinted in Thirlwall, 1997).
Thirlwall, A.P. (1995), *The Economics of Growth and Development: Selected Essays of A.P. Thirlwall*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (1997), *Macroeconomic Issues from a Keynesian Perspective: Selected Essays of A.P. Thirlwall, Volume 2*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (2001), 'The Relation Between the Warranted Growth Rate, the Natural Rate and the Balance of Payments Equilibrium Growth Rate', *Journal of Post Keynesian Economics*, **24**, 81–8.
Thirlwall, A.P. (2002), *The Nature of Economic Growth: An Alternative Framework for Understanding the Performance of Nations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (2005), *Growth and Development with Special Reference to Developing Economies*, Basingstoke: Palgrave Macmillan (8th edition).
Thirlwall, A.P. and Gibson, H. (1992), *Balance of Payments Theory and the United Kingdom Experience*, London: Macmillan (4th edition).
Thirlwall, A.P. and Santos-Paulino, A. (2004), 'The Impact of Trade Liberalisation on Exports, Imports, and the Balance of Payments of Developing Countries', *Economic Journal*, **114**, F50–72.

A.P. Thirlwall: the Kent years

Roger Vickerman

In 2005 the University of Kent celebrated its 40th anniversary and for 39 of those years Tony Thirlwall has been a member of its Department of Economics. As he reaches his own 40th anniversary it is appropriate to reflect on his contribution to the Department of Economics at Kent, as well as to the economics profession. In their Introduction to this volume John McCombie and Philip Arestis have given a more comprehensive account of his contribution to economics, but the two are in many ways inseparable, not least through Tony Thirlwall as a supervisor and mentor of those who started their careers as his graduate students or as junior lecturers in the Department.

Following undergraduate study at Leeds and a Masters programme at Clark, Tony Thirlwall began a PhD in Cambridge, only to be tempted back to a Lectureship at Leeds by his early mentor Arthur Brown, where he finished the PhD. However, in 1966 he made the move south to the newly created University of Kent at Canterbury and, apart from visiting spells at West Virginia, Princeton, Papua New Guinea, Cambridge, Melbourne and La Trobe, he has spent his entire academic career at Kent, rising rapidly from Lecturer to Senior Lecturer to Reader and then Professor of Applied Economics by the age of 35.

The Thirlwall contribution to the reputation of the department is immense. This is not just in the research output: 15 books (and one in press), 12 edited books and approaching 200 chapters and articles in journals, but generations of both undergraduate and postgraduate students have benefited from his lectures, especially on economic development, and the many doctoral students who came specifically to work with him testify to his care and attention as a research supervisor. Those not fortunate enough to study at Kent have nevertheless had the benefit of his clear and precise writing, particularly in the major textbook *Growth and Development: with Special Reference to Developing Economies*, first published as early as 1972 but recently released in its eighth edition, and now with some two-and-a-half times the number of pages as the original. As David Greenaway commented on the seventh edition: 'Few textbooks make it to a seventh edition and even fewer deserve to do so! That is not true of Tony Thirlwall's text, which has been a standard reference for students of Develop-

ment Economics for over thirty years. This excellent new edition will ensure that it remains a standard reference'.

Tony Thirlwall's contribution to economics has come in a number of phases. First came the interest in regional economics and particularly in regional unemployment. This was followed by a more general interest in economic growth (and in particular the role of savings and inflation in the growth process. This led to a specific interest in first of all the role of exports in growth and a specific concern for the way in which the balance of payments exerts a constraint on growth. More recently concern over trade and financial liberalisation has dominated his output. Clearly concern for finding solutions to sets of current economic problems, initially in the UK and then more focused on developing economies, has been a key feature of the way his research interests have developed. But this is not the fleeting interest of the academic dilettante looking for a topical headline or jumping on to the latest conceptual bandwagon rolling through the literature. Throughout all the work there is a consistent view of how the macroeconomy works, a traditional Keynesian view of the importance of aggregate demand strongly influenced by Roy Harrod's and Nicholas Kaldor's insights into the process of growth.

Students, and unsuspecting candidates for lectureships in interviews, have often been caught out by their inability to place what they are doing in the historical literature – the seemingly innocent question 'what is the best book on economics you have read?' followed by 'how did it influence the way you thought about economics?' certainly showed up how much the recipient had really thought about the subject. If they were able to engage in a debate on the chosen volume, then Tony Thirlwall's ability not just to recall a specific reference, but to rehearse an entire line of argument – who said what and in which reference and how they had carried out the analysis demonstrated the mind of the true scholar, but the scholar with an eye for detail and one who really cared about the origin and development of ideas.

Of course, the scholar, as opposed to the more mechanical researcher, cares deeply about his own work and will defend it to the hilt. Littered through the output are examples of debates around the work; comments on pieces written by others and rejoinders to those who have criticised. What features in the published output is only the tip of the iceberg of correspondence – the inveterate letter writer has become more recently the inveterate emailer. This is reinforced on current issues with letters and opinion pieces to newspapers. And where it really matters to him, as for example on the question of UK membership of the Euro, the scholar becomes the pamphleteer and committed member of the 'No' campaign.

This is not the place for a full evaluation of Tony Thirlwall's contribution to economics. That would require a much longer work than this short appreciation, though this volume does some of that through the various contributions which

show how much they have benefited from his own work. But it is fitting to go a little deeper into the main areas of his work to show how Tony's ideas have developed, and how the fundamental belief in the correctness of the Keynesian and Kaldorian insights provides a clear integrity to all the work.

Starting with the regional economics of the early work we see how a real-world problem is analysed using a strong methodological paradigm, harnessing available evidence to produce concrete solutions. The renewed interest in using policy to address the regional problem in the Labour Governments of the 1960s was a strong motivation for this early research. Perhaps ironically, in view of his subsequent move to the UK's southeastern corner, Tony's first published paper was a reply to a piece on 'Restrictions on Expansion in South East England' (Thirlwall, 1965) in which he argued against a view that restrictions were unhelpful in solving problems of regional imbalance. The first key methodological contribution was an important paper on 'Regional Unemployment as a Cyclical Phenomenon' (Thirlwall, 1966a), followed by a more detailed analysis of different types of unemployment (Thirlwall, 1969a). These two papers, supported by a number of others (Harris and Thirlwall, 1968; Thirlwall, 1969c) have become standard references on the regional unemployment problem which have influenced much subsequent work. The culmination of this strand of work comes in the 1970 paper on regional Phillips curves (Thirlwall, 1970), which had an important influence on a young master's student at Strathclyde who later became a Kent colleague from 1979 (Carruth, 1979), and is now also a professor at Kent. However, these were not just academic exercises, they also led to clear policy recommendations and we can observe a pattern in the early output on this topic leading to parallel publications in more general journals explaining the findings to a wider audience (e.g. Thirlwall, 1966b; Thirlwall, 1969b).

In all of these early papers we see the development of a characteristic approach in which a problem is clearly identified and data are used to provide an answer, but all the time with a sound methodological and analytical underpinning. However, by this time the horizon was clearly widening and a number of papers on economic growth are published. These were clearly influenced by the reappraisal of the causes of growth in Kaldor's influential inaugural lecture (Kaldor, 1966), which became a defining feature of much later work. Initially the specific research interest was on savings, inflation and growth, drawn together in a book (Thirlwall, 1974), but the combination of ideas from Kaldor and also from Arthur Lewis led to a stream of publications dealing with, amongst other issues, population growth, labour surplus economies and poverty. Coupled with the success of the *Growth and Development* textbook, this firmly established the reputation of Thirlwall the development economist, with a growing number of research students seeking him out and invitations to get involved in practical work in developing countries.

As well as the development work, a steady stream of papers emerged on various themes, notably an interest in various aspects of the balance of payments, although a number of papers continued to appear on regional unemployment issues. Combining the growth and regional interests, a key paper (Dixon and Thirlwall, 1975a) formalised the ideas of Kaldor (1970) into a formal model structure which provided the basis for considerable further work. This was the most technically developed approach of the attempts by several earlier authors to derive a cumulative causation model of the regional growth process. It was also notable as it represents the first of many papers co-authored with a graduate student, in this case culminating in a book (Dixon and Thirlwall, 1975b). One of the key links in the cumulative causation process is that provided by the Verdoorn relationship, which links productivity growth positively to output growth and helps to overcome the self-balancing effect of diminishing marginal productivity. An English translation of the original 1949 paper by Verdoorn in Italian had been made by Tony Thirlwall and was subsequently published as Thirlwall (1988).

Stemming from this interest in foreign trade and the constraints on growth comes the key work on balance of payments constrained growth and the simple observation in Thirlwall (1979) that has become known as Thirlwall's Law: that the rate of growth is equal to the growth rate of foreign demand multiplied by the ratio of the income elasticity of exports to the income elasticity of demand for imports. This in turn had developed from some joint work with Kent colleague Charles Kennedy building on Harrod's work on trade multipliers (Kennedy and Thirlwall, 1979a, 1979b). Thirlwall's Law led to a significant industry in papers which either supported the result on different data sets or attempted to refute it; either way it firmly established the Thirlwall name in the history of economics. It led also to further work by PhD students (e.g. Thirlwall and Hussain, 1982) and to a highly productive collaboration with John McCombie, a friend and colleague from Cambridge, resulting in two books (McCombie and Thirlwall, 1994, 2004). Interestingly the balance of payments question led to a reappraisal of the relationship with the regional problem (Thirlwall, 1980b). The balance of payments and its constraining impact on growth was not just a key observation in relation to developing countries, the initial observations had been made with respect to the British economy. As well as a whole series of critical papers on approaches to the balance of payments, there was also a textbook (Thirlwall, 1980a), which ran to four editions over the next twelve years with the fourth edition of 1992 co-authored with another Kent colleague, Heather Gibson, now at the Bank of Greece.

Following Nicholas Kaldor's death in 1986, Tony Thirlwall, as Kaldor's literary executor, turned his energy towards the editing of some of Kaldor's later work and two major volumes on Kaldor appeared (Thirlwall, 1987; Targetti and Thirlwall, 1989) as well as a series of biographical papers and notes. A

development of the Kaldorian growth model had also appeared in Thirlwall (1986).

In the 1990s there was a considerable set of parallel interests running, on economic growth and development, on the balance of payments, on Kaldor (and Keynesian economics more generally), much of it generated by the citations (both friendly and critical) to the key earlier papers. The rise of 'new' growth theory posed a challenge to the earlier orthodoxies and this led to a series of papers showing not just that there was less new about these theories than was often claimed but also that often their authors had an inadequate understanding of the history of economic ideas – the scholar over the mere paper writer yet again. This led to further collaboration with younger Kent colleagues such as Khaled Hussain (Hussain and Thirlwall, 2000) and Miguel Léon-Ledesma (Léon-Ledesma and Thirlwall, 2002).

However, new directions were also to take hold as the world economy and some of its impacts on developing countries changed. The era of trade liberalisation and the severe international indebtedness of the poorest countries presented a serious challenge to the concerned scholar. This led to a series of papers, including several written with former PhD students (Thirlwall and Warman, 1994; Santos-Paulino and Thirlwall, 2004; Pacheco-Lopez and Thirlwall, 2004). The desire to collate and summarise this work for the benefit of others is seen in recent shorter treatises (Thirlwall, 2002, 2003). This ability is recognised also in the world-wide appeal of these volumes with Spanish, Portuguese, Japanese and Chinese editions of various books having appeared.

If this relentless activity to probe and test, to seek high standards in the work of others and in students, to seek out collaboration with both distinguished colleagues such as Kaldor and Kennedy, but also to work not just with junior academic colleagues and research students, but also with masters students and more recently undergraduate students on their dissertations (Wells and Thirlwall, 2003), were not enough there is also Thirlwall the editor. First came the series of (initially) biennial Keynes seminars held at Keynes College, at the University of Kent at which eminent scholars, most of whom had known or worked with Keynes, were invited to participate in a seminar on some aspect of Keynes's work and life. Later came the editorial work on Kaldor and other collections based on conferences as well as membership of the Editorial Boards of the *Journal of Development Studies*, the *Journal of Post-Keynesian Economics* and the *African Development Review*. In editorial work the meticulous attention to detail displayed in his own work has been a valuable asset.

What of Tony Thirlwall the teacher and supervisor? Reference has already been made to collaborative work with students. Several generations of PhD students have benefited directly from his advice as supervisor, many others benefited from his advice as Director of the Graduate Programme at Kent from its inception through to 2000. Students from many countries across the world

came to Kent, drawn particularly by the knowledge of *Growth and Development*, but also by word of mouth from one generation to another. They found not only a demanding programme, but also a welcoming environment with an active social dimension culminating in a summer party in the Thirlwall garden. The compulsory badminton tournament on a slope of Himalayan proportions resulted in the winner being presented with a suitably chosen book.

Supervision and teaching has always been meticulous. An admittedly 'old-fashioned' approach with emphasis on a clear exposition of ideas and a discourse around them (Thirlwall the author is clearly recognisable in Thirlwall the lecturer or seminar leader), eschews the use of web-based lecture notes or PowerPoint presentations. But the slightly austere manner does not put students off, they want to take his module on development and they appreciate, even enjoy, the experience. Exposure to Thirlwall the teacher is frequently a basis for the selection of a clearly recognisable dissertation topic by advanced undergraduate or postgraduate students.

As a colleague, and I have worked with Tony Thirlwall for 29 of his 40 years at Kent, life is, as one would expect from such a keen and precise mind, not always easy. The standards which Tony sets himself are ones which he (rightly) expects from other colleagues: hard work, loyalty and above all dedication to the best interests of good students. These traditional virtues frequently do not sit easily alongside the demands of modern universities; student evaluations, peer observation of teaching. The niceties of university budgets and resource allocation models and not least, the Research Assessment Exercise, are all seen to get in the way of the real business of academic life. But that real business is not one of reducing teaching to an absolute minimum in order to concentrate solely on research, but rather one which is student-centred. The lament about modern students is that they do not read enough, do not show the necessary enthusiasm for economics as a discipline and cannot write effectively. There are those who may find the formal scholar remote and not student-friendly (that is, not prepared to spoon-feed information and provide answers on a plate), but any student showing the least spark of interest in a problem will find a generous and willing academic who will read their work with as much attention as an article in a core economics journal.

All in all, whether as a student, colleague or collaborator, working with Tony Thirlwall does not qualify one for a quiet life, but all of us have found that meeting the challenge is rewarding.

REFERENCES

Carruth, A.A. (1979), *Regional Phillips Curves: Some Further Results*, MSc Thesis, University of Strathclyde, Glasgow.

Dixon, R.J. and A.P Thirlwall (1975a), 'A Model of Regional Growth Rate Differences on Kaldorian Lines', *Oxford Economic Papers*, **27**, 201–14.
Dixon, R.J. and A.P Thirlwall (1975b), *Regional Growth and Unemployment in the United Kingdom*, London: Macmillan.
Harris, C.P. and A.P. Thirlwall (1968), 'Interregional Variations in the Cyclical Sensitivity to Unemployment in the UK 1949–64', *Bulletin of the Oxford Institute of Economics and Statistics*, **30**, 55–66.
Hussein, K. and A.P. Thirlwall (2000), 'The *AK* Model of "New" Growth Theory is the Harrod–Domar Growth Equation: Investment and Growth Revisited', *Journal of Post Keynesian Economics*, **22**, 427–35.
Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the United Kingdom*, Cambridge: Cambridge University Press.
Kaldor, N. (1970), 'The Case for Regional Policies', *Scottish Journal of Political Economy*, **17**, 337–47.
Kennedy, C. and A.P. Thirlwall (1979a), 'The Input–Output Formulation of the Foreign Trade Multiplier', *Australian Economic Papers*, **18**, 173–80.
Kennedy, C. and A.P. Thirlwall (1979b), 'Import Penetration, Export Performance and the Harrod Trade Multiplier', *Oxford Economic Papers*, **31**, 303–23.
Léon-Ledesma, M. and A.P. Thirlwall (2002), 'The Endogeneity of the Natural Rate of Growth', *Cambridge Journal of Economics*, **26**, 441–59.
McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance of Payments Constraint*, London: Macmillan.
McCombie, J.S.L. and A.P. Thirlwall (2004), *Essays on Balance of Payments Constrained Growth: Theory and Evidence*, London: Routledge.
Pacheco-Lopez, P. and A.P. Thirlwall (2004), 'Trade Liberalisation in Mexico: Rhetoric and Reality', *Banca Nazionale del Lavoro Quarterly Review*, June.
Santos-Paulino, A. and A.P. Thirlwall (2004), 'The Impact of Trade Liberalisation on Exports, Imports, and the Balance of Payments of Developing Countries', *Economic Journal*, **114**, F50–72.
Targetti, F. and A.P. Thirlwall (eds) (1989), *The Essential Kaldor*, London: Duckworth.
Thirlwall, A.P. (1965), 'Restrictions on Expansion in South East England: A Reply to Mr. Holmans', *Oxford Economic Papers*, **17**, 337–41.
Thirlwall, A.P. (1966a), 'Regional Unemployment as a Cyclical Phenomenon', *Scottish Journal of Political Economy*, **13**, 205–19.
Thirlwall, A.P. (1966b), 'Migration and Regional Unemployment: Some Lessons for Regional Planning', *Westminster Bank Review*, November.
Thirlwall, A.P. (1969a), 'Types of Unemployment: With Special Reference to Non-Demand Deficient Unemployment in Great Britain', *Scottish Journal of Political Economy*, **16**, 20–49.
Thirlwall, A.P. (1969b), 'On the Costs and Benefits of Manpower Policies', *Department of Employment and Productivity Gazette*, November.
Thirlwall, A.P. (1969c), 'Unemployment Compensation as an Automatic Stabiliser', *Bulletin of the Oxford Institute of Economics and Statistics*, **31**, 23–37.
Thirlwall, A.P. (1970), 'Regional Phillips Curves', *Bulletin of the Oxford Institute of Economics and Statistics*, **32**, 19–32.
Thirlwall, A.P. (1974), *Inflation, Saving and Growth in Developing Economies*, London: Macmillan.
Thirlwall, A.P. (1979), 'Balance of Payments Constraint as an Explanation of International Growth Rate Differences', *Banca Nazionale del Lavoro Quarterly Review*, March.

Thirlwall, A.P. (1980a), *Balance of Payments Theory and the United Kingdom Experience*, London: Macmillan.
Thirlwall, A.P. (1980b), 'Regional Problems are Balance of Payments Problems', *Regional Studies*, **14**, 419–26.
Thirlwall, A.P. (1986), 'A General Model of Growth and Development on Kaldorian Lines', *Oxford Economic Papers*, **38**, 199–219.
Thirlwall, A.P. (1987), *Nicholas Kaldor*, London: Harvester Wheatsheaf.
Thirlwall, A.P. (1988), 'Population Growth and Economic Development with Appendix on Factors Governing the Growth of Labour Productivity', (translation of P.J. Verdoorn, 'Fattori che Regalano lo Sviluppo della Produttivita del Lavoro' *L'Industria*, 1949), in D. Ironmonger, J. Perkins and T. Hoa (eds), *National Income and Economic Progress: Essays in Honour of Colin Clark*, London: Macmillan.
Thirlwall, A.P. (2002), *The Nature of Economic Growth: An Alternative Framework for Understanding the Performance of Nations*, Cheltenham UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (2003), *Trade, the Balance of Payments and Exchange Rate Policy in Developing Countries*, Cheltenham UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. and M.N. Hussain (1982), 'The Balance of Payments Constraint, Capital Flows and Growth Rate Differences Between Developing Countries', *Oxford Economic Papers*, **34**, 498–510.
Thirlwall, A.P. and F. Warman (1994), 'Interest Rates, Saving, Investment and Growth in Mexico 1960–90: Tests of the Financial Liberalisation Hypothesis', *Journal of Development Studies*, **30**, 629–49.
Wells, H. and A.P. Thirlwall (2003), 'Testing Kaldor's Growth Laws Across the Countries of Africa', *African Development Review*, **15**, 89–105.

1. The implications of Thirlwall's Law for Africa's development challenges

Mohammed Nureldin Hussain*

In the long run, no country can grow faster than that rate consistent with balance of payments equilibrium on current account unless it can finance ever-growing deficits which, in general, it cannot (A. Thirlwall, 1979).

INTRODUCTION

I first met Professor Thirlwall in 1978 when he was visiting the University of Khartoum as an external examiner for the Department of Economics. At the time I was a newly recruited teaching assistant in the department, and as the youngest staff member, I was pleasantly assigned to look after the social program of the Professor and his family, who had accompanied him to Khartoum. That brief encounter with the Professor was instrumental in resolving a personal quandary: the dilemma of making a choice among the many reputable universities in the United Kingdom and the United States that had accepted me to pursue my postgraduate studies. Impressed by his person and intellect, I joined the University of Kent in 1978 as an MA student. My relationship with Professor Thirlwall grew over time: he was my mentor,[1] my lecturer, my PhD supervisor, and then, Tony, one of my best friends.

It was Tony who first introduced me to the publication business and taught me how to tolerate rejections and to knock on all doors if I had a genuine point to make. With this sprit, we have worked on two models both of which have generated extensive debates. Our work on the supply-side framework for currency devaluation fuelled a long drawn out published and unpublished exchange with International Monetary Fund (IMF) officials.[2] The other work we have done together was on Thirlwall's Law, which is the subject matter of this chapter.

Having worked with Thirlwall's Law in the academic world for some 14 years and in the development practitioners' world for 12 years, I encountered two difficult challenges. The first challenge during my academic career was the usual 'publish or perish'. And in meeting this challenge I always remembered Tony's

advice that it is easy to publish when you swim with the current of the dominant school of thought, but not against it. In the development practitioners' world, in which I am now engaged, the nature of the challenge is different. It is to see that pertinent research finds its way into policymaking instead of being shelved to collect dust. The second challenge is, thus, to avoid what I call 'publish *and* perish'.

I was fortunate enough to be among the few students who read and discussed the manuscript of Thirlwall's Law prior to its publication. I will capitalize on this opportunity to provide a brief documentary of the birth of Thirlwall's Law in an attempt to highlight its deep development implications as well as portray, if only a trace, of the personality and professionalism of the man behind it (section II). The remainder of the chapter will be devoted to discuss the relevance of Thirlwall's Law and its extensions to the contemporary debate on Africa's development issues (section III). These include the question of setting development priorities for poverty reduction in Africa; growth and debt sustainability; globalisation and the income dimension of global competition; and the development effectiveness of foreign aid in relation to the issue of transformation.

THIRLWALL'S LAW: A DOCUMENTARY

It was in the early spring of 1978 when Professor Thirlwall walked into our MA classroom at the University of Kent at Canterbury, England – his flashy smile habitually preceding his 'good mornings'. At the time, he was teaching an MA class of some 18 students, most of them were from overseas and each of them spoke English flavoured by the melody of his/her mother tongue. Upon entering the classroom, Professor Thirlwall had the habit of engaging us in an intelligent and light-hearted chat before addressing the subject matter of the day's lecture. One of his favourite 'chatters' was the pending hazard of Mrs. Thatcher's economic governance – chatter often bejewelled with witty remarks on her belief in supply-side economics. That day there was no chatter. The Professor started the lecturer in silence, distributing first draft copies of a manually typed manuscript that did not enjoy the luxuries of today's word-processing technology. The title of the manuscript was *The Balance of Payments Constraint as an Explanation of International Growth Rate Differences*. A powerful development paradigm was laid before our hands. On first thought, our class, myself included, were more preoccupied with how to decipher the economics of the manuscript for the sake of passing the exams than with appreciating its ramification for development thinking.

Professor Thirlwall embarked on the exposition of his manuscript in his simple but inimitable style. A style that can best be described by what we call in Arabic *the difficult simple*: clear, concise, precise, easy to understand, but

difficult to imitate. It came naturally to him; testimony to his mastery of development economics – theory and practice. He started by saying that the questions of what determined a country's growth rate and why some countries grew faster than others had always been central issues in the study of political economy. He said the theories that attempted to address these questions could be, broadly, grouped into two opposing approaches: the supply orientated and the demand orientated. As usual, he whetted the appetite of the class by engaging them in a series of intriguing questions. He asked whether we believed that supply created its own demand or that demand created its own supply. Some of the young economists in our classroom wanted to be on the winning side with the Professor. They were too green, however, to know which team he supported, despite his humorous remarks on what came to be known as Thatcherism and despite his very apparent resemblance to Lord Keynes in both looks and intellectual orientation. Their answers came head and tails, watching the Professor's face with each throw, lest his sentiments reveal his camp.

Having listened patiently to our answers, the Professor addressed those who genuinely believed in a pre-Keynesian world saying that if supply created its own demand then the rate of output growth would ultimately be explained by the growth rate of the labour force and technical progress. But, 'what determined either of these?', he quizzed the pre-Keynesian believers. He noted the wide use of neoclassical growth theory in development practice notwithstanding, it had very little analytical usefulness in terms of policy implications because it did not explain the true sources of growth. Because I was one of two Africans in the classroom, he turned his lecture towards me. 'Mohammed, let us take economies with fast population growth and abundant labour, such as those in Africa. Could the growth of the labour force alone provide a convincing explanation for the poor economic performance of these countries, their sluggish rates of growth and their inability to utilize the available labour force and other resources?' Before I ventured to answer, he continued, to my great relief, to say that in labour surplus economies it stretched credulity to assume it is an exogenously given supply of labour that determines output in a causal sense. He added that the policy implication of the neoclassical approach would be misleading, in particular, for primary-producing countries like those in Africa. It implied that whatever these countries produced would automatically be sold and lead to higher growth, irrespective of the characteristics of the goods produced and world demand for them.

He nevertheless sought to console, after a fashion, those students who believed in Say's Law. He said that there is, of course, a trivial sense in which the supply-side view of growth was true: that there could be no output without the input of resources – not much by way of consolation, I thought. In order to understand the dynamics of growth, the nature and the extent of demand constraints on the utilization of resources and output growth must be understood, he added.

He postulated that the most binding constraint on growth in an open economy was likely to be the balance of payments.

The Professor started to work out the mathematics of his manuscript. The good old blackboard notwithstanding, the identities and equations of the model were animated, left-handedly, in a manner that competes easily with Bill Gates' PowerPoint facilities. The model contained three basic equations representing the growth of imports, the growth of exports, and a dynamic expression of the overall balance of payments equilibrium. He substituted the first two equations into the third and the model was solved to yield an elaborate expression of the growth rate of real gross domestic product (GDP). When the terms of trade were assumed to be constant the elaborate equation collapsed into an expression containing three symbols: $y = x/\pi$. '*The rate of growth (y) of any developed country in the long run is equal to the growth rate of the volume of its exports (x) divided by its income elasticity of demand for imports (*π*)*', he explained.

Our eyes were fixed on the blackboard, attempting to digest the meaning and internalize the implications of this tri-legged animal. That job was not easy. For the animal distilled volumes of legendary work in economic development, encapsulating all of them in a small-sized anti-underdevelopment pill. The teaching of Engel's law, which implies that the demand for primary goods increases less than proportionally to increases in global income; the Harrod foreign trade multiplier which put forward the idea that the pace of industrial growth could be explained by the principle of the foreign trade multiplier;[3] the Marshall–Lerner condition which implies that a currency devaluation would not be effective unless the devaluation-induced deterioration in the terms of trade is more than offset by the devaluation-induced reduction in the volume of imports and increase in volume of exports; the Hicks super-multiplier[4] which implies that the growth rate of a country is fundamentally governed by the growth rate of its exports; the Prebisch–Singer hypothesis which asserts that a country's international trade that depends on primary goods may inhibit rather than promote economic growth; the Verdoorn–Kaldorian notion that faster growth of output causes a faster growth of productivity, implying the existence of substantial economies of scale;[5] Kaldor's paradox which observed that countries that experienced the greatest decline in their price competitiveness in the post-war period experienced paradoxically an increase in their market share and not a decrease; the literature on export-led growth which asserts that export growth creates a virtuous circle through the link between output growth and productivity growth – all of these doctrines were somehow put into play and epitomized within this small-sized capsule. Not only that but the capsule was sealed by the novel and powerful ingredient of the balance-of-payments constraint: '*in the long run, no country can grow faster than that rate consistent with balance of payments equilibrium on current account unless it can finance ever-growing deficits which, in general, it cannot*'.

The time for class discussion came and all the debate seemed to linger around one basic query: if growth could be explained by a rule which contained two variables only, what was the relevance of the many other socio-economic variables that could also influence the growth process? What about the role of policies and economic management? What about the role of capital, labour and technical progress? The answers of the Professor were convincing to some students, but confusing to many others. In an attempt to relieve our baffled faces he concluded the discussion by saying, in a pleasant fusion of smile and speech: '*Simple laws make good economics*'. And as he was leaving the classroom, his smile turned gradually into a laugh that engulfed his remark: '*if this rule came to be known as Thirlwall's Law, I will retire*'. Less than one year after the publication of the manuscript in 1979 the rule was crowned as '*Thirlwall's Law*'. But, retire? He did not. I suppose genuine philosophers like this man will never retire even if they wanted to.

Professor Thirlwall left our classroom with one query bemusing his mind: the rule might not be of relevance to poor countries that depended heavily on aid inflows. When I enrolled for my PhD under his supervision in 1979, Professor Thirlwall handed me, in one of my frequent and unannounced trips to his office, a hand-written theoretical extension of his balance of payments constraint model that incorporated the effect of capital flows. He invited me to do the empirical work for the extended model by applying it to a group of developing countries. The manuscript rested on my desk for some time before I did the application. Playing back those years, I think the reason for this delay was twofold. First, the extended Thirlwall's model did not fit into the theme of my PhD and I was so keen to make solid progress in my thesis before I could carry what I considered at the time to be an additional load. Secondly, empirical work in those pre-statistical software days was a cumbersome task. To cut a long story short, the regression and other empirical analysis that I could do today in one minute with my PC, used to take us weeks to finish, not to mention the 'long walks' from my study office in Keynes College to the mainframe computer room, with the Siberian wind of Canterbury blowing non-stop over my African head.

At last, I managed to finish the empirical work of the extended Thirlwall's model, and to draft my interpretation of the results. These were incorporated with the theoretical model and the output was a paper titled *The Balance-of-Payments Constraint, Capital Flows and Growth Rate Differences Between Developing Countries*. It took the Professor and myself numerous drafts of the paper, circulated to the University staff and postgraduate students for comments before we were ready to face the publication battlefront – or what we called, in jest, 'the publication mafia'. The extended Thirlwall model was then – after some battles – finally published in *Oxford Economic Papers* in November 1982. Some of the other studies that applied the model for developed and developing

countries include Perraton (1990), Bairam and Dempster (1991), Atesoglu (1993), Andersen (1993), Mulligan (1996), Hussain (1999), Elliott and Rhodd (1999), and León-Ledesma (1999). McCombie and Thirlwall (1994) devote a book of over 600 pages to the critique and defense of the balance of payments model and its superiority to other theories of growth performance. See also McCombie and Thirlwall (1997) for a survey.

THE RELEVANCE OF THIRLWALL'S LAW TO AFRICA'S DEVELOPMENT CHALLENGES

It is now beyond dispute that Africa is the only continent in the world where the number of poor people is rising and where the majority of countries are unlikely able to meet most of the Millennium Development Goals (MDGs). The question of what Africa should do to overcome its development challenges has been addressed by a voluminous body of literature and the agenda for policy action is well known. It is now widely accepted that accelerated GDP growth is the first line of attack in the fight against poverty and Africa would have to achieve a sustained rate of economic growth well above thrice that of population in order to arrest and reverse the spread of poverty. There is also abroad consensus on the policies and strategies that are needed to accelerate economic growth.[6] However, the agenda for action is daunting and no single country in Sub-Saharan Africa can afford to address all the pending challenges, at all fronts simultaneously, and with equal vigour. Additionally the different aspects of the agenda for action are interrelated and interdependent, which makes it hard for African governments, as well as their development partners, to set priorities. This question is pertinent because little attention has been devoted in the literature to the issue of formulating prioritized and long-term development visions for African countries. Previous economic management has been largely externally driven, concentrating on getting macroeconomic fundamentals right, and recent efforts emphasize investing in supply-side factors, particularly human capital development, in the absence of an all-inclusive vision to overcome the basic constraints on growth. It is in this respect that Thirlwall's Law provides a strong theoretical and empirical justification for making the export sector the focal point in setting such a development vision.

Making Exports Promotion Central in the Fight against Poverty

One of the many areas where I find Thirlwall's Law edifying is on the question of how to set priorities for poverty reduction in Africa. The teachings of Thirlwall's Law provide a simple but significant thesis for poor African countries striving for faster growth and for poverty reduction (see Hussain, 2001). While

there is no disagreement that progress in poverty reduction will come through increased investment in physical and human capital that brings about an accelerated and broad-based growth, Thirlwall's Law implies that export expansion plays an important role, not only in making such growth possible, but also sustainable. To explain: fast growth generates a large demand for manufactured capital and consumer goods. In the absence of growing domestic production to satisfy such an increase in demand, the excess demand will spill over into imports, putting pressure on the current account balance. In Thirlwall's Law, this process is put into play by the income elasticity of demand for imports, which is usually very high in African countries because they produce only limited items of consumer goods and virtually no capital goods. Hence efforts to grow faster will generate a faster growth of imports.

If export expansion is not sufficient to meet the import demand associated with faster growth (in addition to servicing foreign debt)[7] the sustainability of fast growth will be threatened, as the heavy burden of foreign indebtedness will eventually close in and capital inflows will eventually dry up. Export growth, whose effect in Thirlwall's model is captured, in the main, by the income elasticity of demand for exports and the growth of world income, is unique as a growth-inducing force from the demand side. This is because it is the only component of demand that provides foreign exchange to pay for the import requirements for growth. In this sense, it allows all other components of demand (consumption, investment, and government expenditure – which all have import-contents) to grow faster in a way that consumption-led growth or investment-led growth does not. In the long run, no country can grow faster than that rate consistent with balance of payments equilibrium on current account unless it can finance ever-growing deficits, which, in general, it cannot (see Thirlwall, 1979 and 1997).

Thus, in contrast to the now dominant supply-side development philosophy that emphasizes investment in physical and human capital *per se,* due regard must be given to the question of how these investments will promote export earnings in their totality and in their composition. A growth strategy that concentrates on expanding investment in human and physical capital without due regard to the 'the foreign exchange productivity of investment' will be short-lived, because the balance-of-payments constraint will eventually put an end to such an expansion, rendering domestic resources, including human capital underutilized and the country heavily indebted (see Hussain, 2001). This pattern typifies the experience of most African countries since independence.

Economic Growth and Debt Sustainability

The persistence of Africa's external gap has led to accumulation of debt and to debt repayments problems. The state of African's heavy indebtedness gave rise,

in Africa's recent past, to numerous debt-relief schemes, the latest of which is the enhanced Heavily Indebted Poor Country (HIPC) Initiative. While the HIPC Initiative is a welcome step towards reducing the debt burden of African countries, its design and implementation have been essentially based on an arbitrary financial criterion, which has not been clearly linked to the nexus of growth and poverty reduction. It should be recalled that under the enhanced framework, the sustainability concept, which applies for most Heavily Indebted Poor Countries (HIPCs) is the ratio of a country's debt to its exports of 150 per cent.[8] This measure dictates that once a HIPC reaches the decision point, it should not allow its NPV debt-to-export ratio to exceed 150 per cent. This measure of sustainability is chosen because of its simplicity and workability. Once the HIPC Initiative entered into its implementation phase, the financially based sustainability criterion encountered serious difficulties. Even some of the models that have been developed subsequent to the emergence of these problems still suffer from the same shortcomings.

The centrality of Thirlwall's Law (and its extensions) to the issues of debt and development is that it brings together, in one analytical apparatus, three aspects important to the question of sustainability: the need to accelerate growth towards achieving socio-economic targets, the need to bridge the financing gap to attain a target growth rate and the need to utilize foreign aid effectively. To illustrate this, it can be argued, in the extreme case, that any poor country can achieve 'debt sustainability' by deliberately curbing its imports, debt repayment, and foreign borrowing to levels that are compatible with its low export earnings. While such a policy can achieve 'sustainability' because it shuts down the sources that can raise the debt-to-export ratio above 150 per cent, the price of such practices is to condemn the country to even lower growth rates and an increased incidence of poverty. This, of course, is contrary to the development objectives of African countries, which aspire to higher economic growth and improved living standards. Clearly, the sustainability of debt must be defined in such a way as to allow African countries to realize their targeted development objectives by expanding their investment beyond the limits permitted by their export earnings – i.e., by borrowing from abroad.

On the other hand, the root cause of the heavy debt of HIPCs is that most of these countries consistently had a current account deficit because their imports were persistently more than their exports, and hence they borrowed from abroad to bridge the financing gap. The persistence of this pattern indicates that the borrowed funds, to bridge the external financing gap, were used in activities that did not alter the pattern of trade to generate sufficient foreign exchange earnings for debt repayment. The search for a comprehensive concept for debt sustainability must also relate to the issue of ensuring effective utilization of the borrowed funds, and in particular how the borrowed funds can be used in improving the capacity of foreign exchange earnings of HIPCs.

In view of this, the present sustainability concept suffers from a number of serious shortcomings. First, there is the problem of economic growth versus debt sustainability. First, the yardstick for determining sustainability has no built-in methodology to cater for the resource needs of these countries to attain the internationally agreed goals such as attaining the MDGs. Secondly, there is no simple formal model to measure the impact of debt relief on growth and poverty reduction. In the same vein, there is no clear development-orientated methodology to assess the impact of external shocks on the gains from the HIPC Initiative and on the HIPCs' long-term debt sustainability. Thirdly, the issue of African countries graduating from aid dependency to self-sustainability has not been clearly catered for either in the design or in the implementation of debt relief. The funds released through HIPC debt relief finance investments in mostly social sectors, which do not generate foreign exchange earnings directly, nor do they change the pattern of production of the countries concerned. Thirlwall's Law provides a useful tool for addressing such questions quantitatively, as well as qualitatively, and in what follows we provide examples of its applicability.

The Impact of Debt Service on Growth and Poverty

There is broad agreement that Africa's debt burden has been a major obstacle to the region's prospects for increased savings and investment; that the continent's debt overhang has inhibited public investment in physical and social infrastructure; that it has also hampered private investment; that it undermined critical investments in health and human resource development – all leading to a reduction in Africa's prospects for accelerated poverty reduction and growth. However, such statements are seldom substantiated by quantifiable empirical evidence. Thirlwall's Law, as extended in Thirlwall and Hussain (1982) and in Elliott and Rhodd (1999), provides a convenient methodology to estimate the impact of the debt burden on the growth rate of individual African countries, and the impact of debt on poverty could also be estimated in conjunction with the poverty elasticity of growth. In Hussain and Gunter (2005), this methodology was applied to 44 African countries, 22 of which had reached the enhanced HIPC decision point by May 2005. The estimates covered the period 1985–1999 and the results are summarized in Table 1.1. The table shows the effect of debt service on growth separately in column (A), while all the other effects of the model are grouped in column (B).[9]

It can be seen from Table 1.1 that the growth rates estimated by the model are very close to the actual growth rates. This validates the model as a good predictor of the growth experience of African countries.[10] The application of the model also reveals that debt service had a large negative effect on the growth rate of African countries. On average, debt repayment reduced the growth rate

*Table 1.1 The effect of debt service on real GDP growth and poverty in Africa (per annum, 1985–1999 averages)**

Country Name	Predicted Effect of Debt Service on Growth (A)	Predicted Other Effects on Growth (B)	Predicted Growth Rate (C = A+B)	Actual Growth Rate	Poverty Elasticity of Growth (D)	Predicted Effect of Debt Service on Poverty (E = C*D)
I. HIPCs (that reached the enhanced DP by Dec. 2000)						
BENIN	–1.3%	5.2%	3.9%	3.4%	–1.12	1.4%
BURKINA FASO	–1.9%	9.9%	8.0%	6.0%	–0.69	1.3%
CAMEROON	–1.6%	0.9%	–0.7%	0.5%	–0.38	0.6%
GAMBIA, THE	–0.8%	1.2%	0.4%	3.8%	–0.88	0.7%
GUINEA	–7.1%	8.4%	1.3%	4.3%	–0.38	2.7%
GUINEA-BISSAU	–2.1%	1.0%	–1.1%	1.7%	–0.38	0.8%
MADAGASCAR	–3.4%	4.9%	1.5%	1.9%	–0.80	2.7%
MALAWI	–4.3%	6.3%	2.0%	3.5%	–0.52	2.2%
MALI	–2.4%	7.9%	5.5%	4.8%	–0.56	1.4%
MAURITANIA	–1.4%	6.5%	5.1%	3.4%	–1.57	2.2%
MOZAMBIQUE	–6.7%	9.8%	3.1%	5.6%	–0.52	3.5%
NIGER	–0.9%	5.4%	4.5%	2.7%	–0.56	0.5%
RWANDA	–3.8%	5.7%	2.0%	1.8%	–0.38	1.4%
SAO TOME & PRINCIPE	–4.0%	3.8%	–0.3%	1.2%	–0.38	1.5%
SENEGAL	–3.3%	3.7%	0.4%	3.0%	–1.96	6.4%
TANZANIA	–7.9%	11.9%	4.0%	3.8%	–0.81	6.6%
UGANDA	–3.6%	6.8%	3.2%	5.3%	–1.04	3.8%
ZAMBIA	–10.1%	10.5%	0.4%	0.3%	–0.42	4.2%
Average (PPP weighted)	**–4.1%**	**6.5%**	**2.4%**	**3.3%**	**–0.75**	**2.8%**
II. HIPCS (that reached enhanced DP by April 2004)						
CHAD	–0.3%	2.8%	2.5%	1.0%	–0.38	0.1%
ETHIOPIA	–8.8%	12.0%	3.1%	3.3%	–0.67	5.9%
GHANA	–6.7%	10.1%	3.4%	4.3%	–2.40	16.0%
SIERRA LEONE	–10.6%	7.2%	–3.4%	–1.4%	–1.12	11.9%
Average (PPP weighted)	**–7.4%**	**10.3%**	**2.8%**	**3.3%**	**–1.36**	**9.8%**

III. HIPCS (that have not yet reached enhanced DP)						
BURUNDI	–2.1%	4.5%	2.4%	1.8%	–0.38	0.8%
COMOROS	–0.2%	3.3%	3.1%	2.7%	–0.94	0.2%
CONGO, REPUBLIC OF	–1.7%	5.4%	3.7%	3.4%	–0.38	0.6%
COTE D IVOIRE	–5.6%	7.4%	1.8%	3.1%	–1.79	10.1%
TOGO	–0.7%	3.4%	2.7%	2.8%	–1.12	0.8%
Average (PPP weighted)	**–4.0%**	**6.1%**	**2.2%**	**2.9%**	**–1.37**	**6.5%**
IV. Other African countries						
ALGERIA	–8.1%	11.0%	2.9%	2.4%	–1.57	12.7%
BOTSWANA	–1.2%	8.4%	7.2%	7.2%	–0.52	0.6%
CAPE VERDE	–1.3%	3.9%	2.6%	2.1%	–1.12	1.5%
EGYPT	–3.8%	8.6%	4.9%	4.0%	–1.57	5.9%
EQUATORIAL GUINEA	–6.1%	16.5%	10.4%	14.0%	–0.38	2.3%
GABON	–0.7%	2.1%	1.5%	0.9%	–0.38	0.2%
KENYA (sustainable HIPC)	–2.6%	6.4%	3.7%	0.9%	–1.40	3.7%
LESOTHO	–0.9%	6.2%	5.3%	4.8%	–0.52	0.5%
MAURITIUS	–1.8%	10.7%	8.9%	8.8%	–0.94	1.7%
MOROCCO	–4.7%	8.7%	4.0%	3.4%	–1.57	7.4%
NAMIBIA	–0.6%	5.4%	4.7%	4.4%	–0.52	0.3%
NIGERIA	–10.4%	12.2%	1.8%	–1.7%	–1.16	12.0%
SEYCHELLES	–1.9%	9.3%	7.4%	4.3%	–0.94	1.8%
SOUTH AFRICA	–1.8%	2.7%	0.9%	0.6%	–0.52	0.9%
SWAZILAND	–1.4%	8.1%	6.7%	7.2%	–0.61	0.9%
TUNISIA	–2.0%	5.5%	3.5%	3.5%	–1.57	3.1%
ZIMBABWE	–11.9%	15.1%	3.2%	0.5%	–0.52	6.1%
Average (PPP weighted)	**–4.4%**	**7.1%**	**2.7%**	**1.8%**	**–1.07**	**5.4%**
14 Afr. Compl. Point HIPCs (PPP w.)	**–5.6%**	**8.8%**	**3.3%**	**3.8%**	**–1.08**	**5.9%**
All African HIPCs (PPP w.)	**–5.0%**	**7.5%**	**2.5%**	**3.2%**	**–1.00**	**5.3%**
Africa (PPP weighted)	**–4.5%**	**7.2%**	**2.6%**	**2.1%**	**–1.06**	**5.4%**

Note: * Angola, the Central African Republic (CAR), the Democratic Republic of Congo (DRC), Djibouti, Eritrea, Liberia, Libya, Somalia, and Sudan are excluded due to lack of data of these nine excluded countries, the DRC reached the enhanced decision point in July 2003; Angola, Djibouti, Eritrea, and Libya are considered sustainable without requiring HIPC debt relief; the CAR, Liberia, Somalia, and the Sudan remain to be considered under the HIPC Initiative.

Source: Hussain and Gunter (2005).

of HIPCs by 5.0 percentage points per annum between 1985–1999. As would be expected, the reduction of the rate of growth as a result of debt repayment was smaller in the case of non-HIPCs, amounting to an average –4.4 per cent per annum, and it averaged –4.5 per cent per annum for all 44 countries in the sample. Interpreting these results differently, it could be said that in the absence of debt service, the growth rate of HIPCs would have averaged 8.2 per cent per annum instead of 3.2 per cent per annum between 1985–1999. The respective figures are 6.2 per cent for non-HIPCs and 6.6 per cent for the whole sample of African countries.

Using the poverty elasticity of growth (second last column of Table 1.1) it can be easily shown that the reduction in the pace of economic growth resulting from the leakage of resources through debt services has caused large increases in poverty (as measured by the proportion of people living under the poverty line of one dollar a day). According to our estimates, the negative impact of debt service on growth had caused poverty to increase in HIPCs by an average of 5.3 per cent annually during the period of estimation. The increase in poverty as a result of debt services averaged 5.4 per cent annually for the whole sample of African countries.

The Impact of the HIPC Initiative on Growth and on Poverty Reduction

The HIPC framework has been designed to reduce the negative impact of the debt burden through directly reducing the stock of debt and hence debt service. Through linking debt relief to economic and social reforms, the framework has also been expected to reduce poverty by offering the possibility of releasing significantly larger resources for investments in such areas as education and health. In the extended Thirlwall's model, the direct impact of debt relief on growth works through the reduction in the rate of change of debt services. In Hussain and Gunter (2005), the net effect of HIPC debt relief on growth has been measured using the before and after method (Table 1.2) for the 18 African HIPCs that have reached their enhanced HIPC decision point by end-2000. It can be observed from Table 1.2 that in the period prior to the implementation of HIPC debt relief, debt service had a large negative impact that averaged –4.1 percentage points per annum. In the period after reaching the enhanced decision point (i.e., when receiving HIPC debt relief) the effect of debt services on growth fell to –1.2 percentage points. Taking the difference between the two, this indicates that HIPC debt relief had a net positive effect on growth that amounted to an average of 2.9 per cent per annum between 2001 and 2003. In other words, everything remaining the same, the growth rate in HIPCs would have increased by an average of 2.9 per cent as a result of the implementation of the HIPC Initiative. This also implies a reduction in poverty by an average of 2.2 per cent per annum between 2001 and 2003. However, in the context of the extended

Table 1.2 Net increase in real GDP growth and net reduction in poverty due to the HIPC Initiative in Africa (2001–03 averages)

Country Name	Effect of Debt Service on Growth		Net Increase in Real GDP Growth due to the HIPC Initiative	Effect of Debt Service on Povery		Net Reduction in Poverty due to the HIPC Initiative
	Before the HIPC Initiative (1985–99)	After the HIPC Initiative (2001–03)		Before the HIPC Initiative (1985–99)	After the HIPC Initiative (2001–03)	
BENIN	–1.3%	–0.4%	0.9%	1.4%	0.5%	1.0%
BURKINA FASO	–1.9%	–0.9%	1.0%	1.3%	0.6%	0.7%
CAMEROON	–1.6%	–0.4%	1.1%	0.6%	0.2%	0.4%
GAMBIA, THE	–0.8%	–0.4%	0.4%	0.7%	0.4%	0.4%
GUINEA	–7.1%	–1.6%	5.5%	2.7%	0.6%	2.1%
GUINEA-BISSAU	–2.1%	–0.7%	1.4%	0.8%	0.3%	0.5%
MADAGASCAR	–3.4%	–0.6%	2.8%	2.7%	0.5%	2.2%
MALAWI	–4.3%	–2.3%	2.1%	2.2%	1.2%	1.1%
MALI	–2.4%	–0.8%	1.6%	1.4%	0.4%	0.9%
MAURITANIA	–1.4%	–0.6%	0.8%	2.2%	0.9%	1.3%
MOZAMBIQUE	–6.7%	–2.2%	4.5%	3.5%	1.1%	2.3%
NIGER	–0.9%	–0.3%	0.6%	0.5%	0.2%	0.3%
RWANDA	–3.8%	–2.1%	1.7%	1.4%	0.8%	0.7%
SAO TOME & PRINCIPE	–4.0%	–3.1%	1.0%	1.5%	1.2%	0.4%
SENEGAL	–3.3%	–0.7%	2.6%	6.4%	1.3%	5.1%
TANZANIA	–7.9%	–1.9%	6.0%	6.4%	1.5%	4.9%
UGANDA	–3.6%	–0.7%	3.0%	3.8%	0.7%	3.1%
ZAMBIA	–10.1%	–4.2%	5.9%	4.2%	1.8%	2.5%
Average (PPP weighted)	–4.1%	–1.2%	2.9%	2.8%	0.8%	2.2%

Source: Hussain and Gunter (2005).

Thirlwall's Law these are once-and-for-all effects and they can be eroded by such external shocks as deteriorations in the terms of trade for HIPCs. Also the effects of the HIPC Initiative are once-and-for-all effects and are not expected to continue unless debt service continues to decline. This appears unlikely, particularly if these countries attempt to bridge the large financing gap to secure the resources required by African countries to accelerate their growth rate to the levels compatible with the attainment of the MDGs.

The Financing Gap and Development Effectiveness[11]

The introduction of the MDGs has brought to the forefront the question of how to measure the resource gap required by poor countries to accelerate their growth rate to the levels required to attain these objectives. The dilemma faced by the development community in this regard is the absence of a credible model to estimate such a gap. The Harrod–Domar model, which was the traditional methodology used to estimate the financing gap, had been largely discredited. It might be recalled that in 1997, William Easterly authored an intriguing paper titled *The Ghost of Financing Gap: How the Harrod–Domar Model Still Haunts Development Economics.* His investigation shows that the Harrod–Domar model fails miserably in explaining the relation between investment and growth. He reveals that the Harrod–Domar model, which supposedly died 40 years ago, is still used by leading international and regional organizations. Indeed, more than 90 per cent of the country economists in the World Bank still use the model to make growth and resource requirements projections, and to provide advice on economic policy. The demise of the Harrod–Domar model has created a true dilemma for development practitioners: the 'absence' of a credible alternative.

Hussain (2001) has shown that the extended version of Thirlwall's Law could be used as an alternative approach to the discredited Harrod–Domar model. He has shown that the model based on Thirlwall's Law can be used to predict growth, to measure the financing gap, to formulate policy advice, and to provide pointers for measuring the development effectiveness of foreign development assistance. Comparing the Harrod–Domar model with the one based on the extended Thirlwall's Law, it has been shown that the former is susceptible to the misinterpretation that all capital inflows (mainly foreign assistance in the case of Africa) required to fill the financing gap will be invested. This is because the Harrod–Domar model focuses on the saving-investment relationship in a planned sense (*ex ante*). In practice, once the foreign exchange to bridge the 'planned' resource gap is obtained, there is no guarantee that all of it will finance investment.

Stemming from Thirlwall's Law it has been argued that it is the faulty assumption that all aid is intended for investment that might have partly contributed to the negative assessment of aid effectiveness, and hence, to the diminished

public support in donor countries for aid programs. It has been revealed that in many cases, capital inflows to bridge the resource gap, particularly under program aid, were predestined, by design, for consumption and not for investment. In the post-oil crisis era, for instance, a large proportion of concessional resources from multilateral sources have been (and continue to be) in the form of balance of payments support under reform programs. Such concessional flows are essentially intended to cushion the implementation of reforms and not to increase the level of investment. In the case of trade liberalization, for example, concessional resources are meant to finance the 'liberalized' imports without clear distinctions between consumer and capital goods. They might also be used to compensate the government for the loss of trade duties irrespective of whether the implied government expenditure is destined for consumption or investment. In this case, the 'effectiveness of aid' should probably be assessed in terms of its ability to increase imports of consumer goods and not investment.

Additionally it has been demonstrated that Thirlwall's Law has the advantage of accommodating the terms-of-trade effect on the estimates of foreign exchange requirements. There is the long-standing argument that whatever African countries get in the form of international assistance, they might lose a greater amount in the form of deterioration in the terms of trade. Thirlwall's approach, by including a separate variable for the terms of trade, permits the practitioner to gauge their effects on resource requirements in conjunction with the price elasticities for imports and exports. The approach also gives a measure of how inflation, exchange rate policies, and export and import volumes affect the estimates of the financing gap.

The Thirlwall extended model has been used to measure the financing gap in a sample of 24 African countries, against the objective of reducing the number of people living in poverty by 4 per cent per annum to attain the MDG on poverty, as a requirement for achieving the international development goal of reducing poverty by 50 per cent by the year 2015. The results showed that given the terms of trade, export volume and initial imbalance in the current account, virtually all the countries in the sample would require large inflows of foreign resources to achieve this target. If export performance were not improved relative to imports, the financing gap would increase over time, snowballing to very large amounts that would jeopardize present sustainability measures. For the 24 countries in the sample, the financing gap is estimated at a total of US $44.8 billion (or 21 per cent of projected GDP) per annum in the first five planning years alone. The cumulative amount of resources for all the 24 countries over the first five planning years reaches about US $224 billion, with serious implications for the accumulation of debt and debt repayments. This takes us to the issue of the development effectiveness of foreign aid.

From the extended Thirlwall's Law it can be deduced that foreign aid can contribute to higher growth rates because it finances the excess of imports over

exports. However such a dependency on aid to support higher growth rates will continue unless the production structure and the pattern of trade in the recipient country are changed towards the production of more attractive goods and services. If the production pattern is not changed, the country will slide back to a lower growth path when foreign aid dries up. Thus, if the ultimate goal of foreign assistance is to help poor countries graduate to a self-sustaining growth path, the model suggests two broad pointers for measuring the long-term development effectiveness of foreign aid. That is, the development effectiveness of foreign aid should be measured in terms of its ability to promote export growth relative to that of imports in the recipient country and/or its ability to create the conditions that will act as a catalyst to attract private capital. Even in this latter case, Thirlwall's Law implies that the development effectiveness of private capital inflows must also be measured in terms of their contribution to expanding export earnings relative to imports.

The Income Dimension of Global Competition

Globalisation, which is an irreversible force, driven essentially by the revolution in information technology and liberalization of trade and capital, is one of the challenges facing Africa's development. Recent studies have found that trade is the single most powerful vehicle through which the economic gains from globalisation are distributed between nations.[12] This finding is not new. Historically, trade has acted as an important engine of growth for countries at widely differing stages of development, not only by contributing to a more efficient allocation of resources within countries, but also by transmitting growth from one part of the world to another. Globalisation has reinforced this central role of trade in the process of growth and development.

However, not all countries necessarily share equally in the globalisation-induced growth of trade or in its benefits. One of the fundamental teachings of Thirlwall's Law is that the distribution of the gains from trade rests crucially on a country's income elasticities of demand for exports relative to that of imports. It is ironic that despite the importance of these two income elasticities for Africa's trade performance, and hence its socio-economic development, they rarely figure in the literature and discussions of the leading development institutions (Hussain, 2001). Even elementary theory tells us that demand for goods and services stands basically on two legs, namely, relative prices and income (global income in the case of the demand for exports). It is indeed astounding that the policies formulated by the leading institutions in the area of trade continue to be one-footed – concentrating on price-competition and overlooking the role of global income as crucial elements in the competitiveness equation. Policies to expand exports of African countries were mainly confined to improving price competitiveness through currency devaluation, with the percentage devaluation usually measured

to make traditional exports price-competitive.[13] This policy, though it might be useful in the short run if certain conditions were satisfied,[14] has two major 'strategic' shortcomings. It enforces the dependency of African countries on primary exports – the very reason for Africa's structural trade weaknesses in the first place – and it ignores the all important role of non-price competition.

Africa has continued to fare quite poorly in its international trade performance over the last two decades. Its share in world exports has declined to 2 per cent in 2002, compared to 3 per cent in 1990, and to 6 per cent in 1980.[15] The loss of Africa's market share over the last two decades or so is a clear indication that the policy of expanding its exports through improving price-competitiveness was not successful. It has been shown that Africa's price competition explains only a small proportion of the changes in its market shares; non-price factors appear to be of much greater importance in determining the success, or otherwise, of a country in the international market.[16] And Africa is not unique in this. There is overwhelming evidence from other continents that price competitiveness is quantitatively unimportant in determining changes in international market share.[17] For Africa to reap the benefits of globalisation and improve its trade and growth performance increased attention would, thus, need to be devoted to improving non-price competition, which is captured by the income elasticity of demand for exports.[18]

Non-price competition encompasses all those factors, other than price, that affect a country's market share. These factors are normally grouped into two broad categories: the first is related to the act of selling or marketing; and the second is related to the characteristics of the product including its level of sophistication. Improving these non-price aspects will tend to increase the income elasticity of demand for Africa's exports.[19] The first aspect of non-price competition, which relates to the act of marketing, is determined by such factors as infrastructure; packaging; communication and foreign contacts; export processing services; quality controls and hygiene; the speed of delivery of products; the provision of export-servicing facilities in ports and airports; and, the existence of effective trade promotion organizations. Marketing depends on the improvement of packaging; foreign contacts; the speed of delivery of products and the provision of after-sale services in the case of durable goods. The provision of these factors in Africa falls short both in comparison with other developing countries as well as in relation to the needs of the African peoples. Infrastructure in Africa including roads, ports, airports and export-servicing facilities is generally much lower than that of Asia and Latin America. It is not surprising, therefore, that African countries have even failed to boost their competitive edge in the field of primary commodities, despite the sweeping currency devaluations, resulting in large losses in market shares.[20]

The second aspect of non-price competition relates to the characteristics of the products produced and exported and their level of sophistication. This gave

rise to what I called in a previous paper[21] the '*income dimension of global competition*'. This latter term is essentially a derivative of the works of McCombie in his defence of Thirlwall's Law.[22] It finds its roots in the 'inverse of Engel's law', which implies that the demand for knowledge-intensive goods increases more than proportionally to increases in global incomes. Perhaps, the most demanding challenge facing African countries in today's globalized world is in the domain of 'the income-dimension of global competition'.

Table 1.3 shows an index of competitiveness as viewed from the income dimension perspective. The index is constructed by estimating global income elasticities for the major commodity groups traded in the international market.[23] Emphasis has been placed on global elasticities (as opposed to country income elasticities), in order to net out the effects of country-specific marking aspects such as transport that influence the magnitude of country-specific income elasticities, thereby concentrating on the pure income dimension. Internationally traded goods have been arranged in the table according to the magnitude of their global income elasticity from the largest to the smallest.

It can be seen that the category of machinery and transport equipment is the most income-attractive category with a global income elasticity of 2.30. This indicates that a 1 per cent increase in global GDP will cause the exports of this category to increase by 2.30 per cent. This is followed by chemical products (2.14) and manufactured goods (2.13). The least income-attractive category is agricultural raw materials with a global income elasticity of 0.74 followed by iron and steel with a global income elasticity of 0.91. The index of the pure income-competitiveness has then been computed by weighting these elasticities by the global shares of countries (or continents) in the respective categories of the internationally traded goods. It is evident from Table 1.3 that Africa is trading in commodities which are highly unattractive to global income. Africa's index of global income competition is only 0.92 compared to 1.76 for Asia, 1.87 for the European Union (EU), 2.02 for the United States of America (USA), and 2.1 for Japan.

The result of this low income-competition is that the market size for Africa's traditional exports is shrinking relative to the global export market. This implies that as world income grows, a smaller proportion of this income growth will be devoted to the purchase of these commodities. The market share of most primary products – such as those exported by African countries – in the global export market has shrunk by an average of 70 per cent between 1970 and the late 1990s.[24] Also, while many developing countries, particularly those in South and South East Asia, have experienced a decrease in the share of primary commodities to total exports, bringing their level of commodity dependence close to that of developed countries, a total of 39 out of 47 African countries are dependent on just 2 primary commodities for over 50 per cent of their export earnings.

Table 1.3 Estimates of an index of the income dimension of global competition[a]

Real World Exports by category	Global Income Elasticity	Shares in Global Trade of Each Category				
		Africa	Asia	USA	EU	Japan
1. Machinery and transport equipment	2.30	0.02	0.11	0.281	0.192	0.355
2. Chemical products	2.14	0.01	0.06	0.066	0.056	0.039
3. Manufactured goods	2.13	0.02	0.353	0.46	0.411	0.492
4. Other manufactured goods	1.38	0.03	0.211	0.113	0.164	0.098
5. Primary commodities, including fuel	1.35	0.03	0.095	0.011	0.025	0.002
6. All food items	1.26	0.05	0.105	0.045	0.063	0.004
7. Ores and metals	1.16	0.1	0.053	0.011	0.069	0.007
8. Agricultural raw materials	0.74	0.75	0.013	0.013	0.02	0.003
Weighted Global Income Elasticities		**0.918**	**1.756**	**2.017**	**1.871**	**2.101**

Note: [a] All elasticities are significant at the 95% level of confidence.

Source: Author's estimates, using data from the Statistics Division of the African Development Bank.

Competition in the global market arena among the leading countries is now conducted in terms of 'strategic visionary planning': the quest to acquire competitive advantage in new industries where future world demand will be high, technological progress will be rapid and labour productivity will rise fast. Paying more attention to the income dimension of global competition means that African countries should strive to build dynamic export sectors and acquire new competitive advantages in products that are more attractive to world income.

CONCLUSION AND POLICY IMPLICATIONS

I have attempted in this chapter to demonstrate that Thirlwall's Law (and its extensions) have a multitude of analytical capabilities and pertinent policy implications for African economies in their struggle to accelerate economic growth and reduce poverty. For one thing, the Law brings together, in one analytical apparatus, the factors which are central to Africa's economic prospects: the important nexus of economic growth, the terms of trade, the volumes of trade, capital flows, and foreign debt. Reports on Africa's development challenges invariably continue to repeat the simplistic view that Africa is poor because its income is low, and that its income is low because its investment is low, and that its investment is low because its saving is low. This simplistic analysis often leads to the conclusion that an injection of foreign aid would, therefore, be needed to supplement domestic savings to boost investment and break up the shackles of this vicious circle. This has been the state of the art and the practice of the development community since the independence of African countries. However, there is little evidence, so far, that the shackles of the vicious circle have been broken. The incidence of poverty is yet to be reversed and the injection of foreign aid has often resulted in adding to the debt burden flowed by efforts at debt relief.

Thirlwall's Law represents a complete departure from this simplistic supply-side view. It asserts instead that to break up the vicious cycle of low income and rising poverty, the main focus of economic management should be on changing the pattern of exports. This is because the current account position of a country is the main constraint on economic growth, because it imposes a limit on the demand to which supply can adapt. Seen through the lenses of Thirlwall's Law, the injection of foreign capital will surely lead to higher growth rates, but given the production pattern of African countries such higher growth rates will not be sustainable, and the countries would have to slide back to lower growth rates as debt accumulates (through the leakage of demand into imports that are not matched by export generation), as capital inflows slow down (because of decreased creditworthiness), and as debt repayments start to impose a heavy

burden on domestic economies (through the leakages of domestic resources into debt repayment).

Seen through the lenses of Thirlwall's Law, Africa's debt burden has been, in large part, the result of the fact that the funds borrowed from abroad over the years were not effectively used to relax the balance of payment constraint on growth by raising the income elasticities of demand for African exports relative to those of imports. Putting it more directly, the injection of foreign capital, concessional or otherwise, was not used to effectively transform the pattern of production of African countries, which still depend on exports of primary goods. It is for this very reason that the present debt sustainability concept is incompatible with Africa's quest for accelerated growth and poverty reduction. Given Africa's export pattern and present levels of capital inflows to Africa, the sustainability of foreign debt can only be achieved at growth rates which are, by far, lower than those required to reduce poverty and achieve the MDGs.

Should the ultimate goal of foreign assistance be to help poor countries graduate to a self-sustaining growth path that depends on its own resources and private capital flows, the development effectiveness of foreign aid should be measured in terms of foreign exchange earnings. As it has been concluded in Hussain (2001), the basic lesson for development practice is that when foreign exchange is the binding constraint, which is the case for most African countries, overall rates of return on aggregate investment ought to be measured in terms of foreign exchange earnings. This does not mean that all foreign assistance must finance activities in the export sector. Rather, it means that African governments must make sure that foreign capital in their aggregates must work to transform the export pattern and generate foreign exchange earnings well and above those required to repay the debt. Those involved in development practice should, therefore concentrate more on innovative variables that reflect the foreign exchange productivity of aggregate investment.

The policy implications of Thirlwall's Law for development planning are that governments that want to wage a true war against poverty through accelerated and sustainable growth would need to adopt a long-term visionary strategy of socio-economic development that is firmly anchored on the promotion of income-attractive importable and exportable goods. Regulatory, investment and trade policies, the pursuit of good governance, institutional building, infrastructure development, and human capital formation, all must be designed and implemented with this fundamental objective in mind.

A detailed discussion of policies for effective economic transformation is beyond the space available for this chapter. However, there are certain planks that can be outlined in brief. While the transformation strategy should, by necessity, be country-specific, there are common guideposts. First, there must be a developmentally oriented and stable government that adopts a long-term visionary plan with the aim of carving out, among other objectives, new export

niches. Such a vision must have wide support from the public, the private sector and political parties. Experience has shown that it is not necessary for a country to develop an integrated national industry in order to initiate a successful transformation process. Greater and higher benefits accrue to those economies that are able to carve out niches for themselves in the global market arena. The export sector is envisaged as taking up the challenge of processing and manufacturing, aiming to advance certain manufacturers into a renowned market niche, encouraging and hosting foreign investment, and acting as the vehicle which pulls the overall economy to higher growth rates and as the window through which the country would acquire new technological knowledge and managerial knowhow.

Special export processing zones (EPZs) would need to be used intensively. There are two main justifications, in the context of African economies, for using such zones. First, that the process of transformation through industrialization requires huge investment in infrastructure and other production-servicing facilities that most African countries may not be able to provide country-wide. The available resources can be devoted to the provision of these facilities within a small confined area, the EPZ. Second, the effective application of countrywide export incentives such as exemption or rebate systems, are more easily monitored in a confined EPZ. Furthermore, the education system, particularly higher and technical education must be harmonized with the specific needs of the private sector, particularly those that are involved in the development of new export niches. In a nutshell, transformation efforts must be anchored on the realization that comparative advantage is no longer static and nature-given as viewed by classical trade theory. It is rather dynamic, man-made and can be deliberately acquired. Africa's quest for fighting poverty will hinge crucially on efforts to build dynamic export sectors and acquire new comparative advantages in exports where world demand will be high, technological progress will be rapid, and labour productivity will rise fast.

NOTES

* Manager, Research Division, African Development Bank, Tunis, Tunisia, until his untimely death on July 15, 2005. The views expressed in this chapter are not to be taken as those of the African Development Bank. The author is grateful to Kouako Koua Louis for some data collection and initial computation and to Rhoda Bangurah for copy-editing. He is also thankful to Dr. Zeinab El Bakri, Dr. Bernhard G. Gunter, and Dr. Barfour Osei for their useful comments and to his wife Tunna for support.

1. Professor Thirlwall had always been very keen to groom his PhD students as 'typical English professors' in conduct as well as intellect. I still remember when he called me to his office one day in the summer of 1979 to inquire angrily whether I was a postgraduate student or a guitarist. The reason was that I was dressed in a fancy white summer jacket with some fancy script printed on its back that read: 'Star Tuck'. My supposition was that the Professor was upset because he confused the 'T' for another letter.

2. For this exchange see Hussain and Thirlwall (1984) and Hussain and Thirlwall (1986).
3. See Kaldor (1978).
4. See McCombie (1985).
5. See Kaldor (1975).
6. These include first, deepening economic reforms, and promoting the development of the private sector so as to strengthen the sustainability of investment. Second, building human capital through efficient delivery of social services – principally health, education and nutrition. Third, promoting regional integration, which provides opportunities to pool resources for investment and to enhance the efficiency of production by taking advantage of the economies of scale that larger markets make possible. Fourth, addressing the long-term foundations of development – notably gender mainstreaming, environmental sustainability and the pursuit of good governance. And with regard to the question of good governance, the imperative of resolving societal conflicts and building the institutions that foster harmony among our peoples and facilitate orderly settlement of disputes without resort to violence must be underscored.
7. The effects of debt servicing in the context of the Thirlwall's Law are discussed in Elliott and Rhodd (1999).
8. For those HIPCs with very open economies where exclusive reliance on external indicators may not adequately reflect the fiscal burden of external debt, a net present value (NPV) debt-to-export target below 150 per cent can be recommended if the country concerned meets two criteria at the decision point: an export-to-GDP ratio of at least 30 per cent and a minimum threshold of fiscal revenue in relation to GDP of 15 per cent. For these countries, the NPV debt-to-export target will be set at a level that achieves a 250 per cent of the NPV debt-to-revenue ratio at the decision point.
9. The impact of debt service of growth is measured as:

$$g_d = \frac{-\frac{D}{T}(d+e)}{\pi}$$

   where D is total debt service repayments, principal plus interest, T is total payments equal to total imports plus total debt service repayments, d is the rate of change in total debt services, e is the exchange rate defined as the domestic price of foreign currency and π is the income elasticity of demand for imports.
10. For a discussion on the methods of testing Thirlwall's Law with its different versions see McCombie and Thirlwall (1994) and Hussain (1999).
11. This section draws on Hussain (2001).
12. See African Development Bank (2003 and 2004).
13. See the exchange between Hussain and Thirlwall (1986) and Nashashibi and Clawson (1986).
14. The success of currency devaluation in improving a country's external account is not automatic. It depends on satisfying certain conditions. If viewed from the demand-side it would have to satisfy the famous Marshall-Lerner condition, if viewed from the supply-side it has to satisfy the Hussain-Thirlwall condition (see Hussain and Thirlwall, 1984).
15. See African Development Bank (2004).
16. See Hussain (1998).
17. For a discussion of this issue see McCombie and Thirlwall (1994).
18. See McCombie (1992).
19. For the relationship between non-price factors and the income elasticity of demand for imports see McCombie and Thirlwall (1994).
20. For instance, it has been estimated (see African Development Bank, 1995) that over the last 25 years (1970–1994) the region's market shares in cocoa beans fell from 80 per cent to 67 per cent; in coffee from 26 per cent to 15 per cent; in cotton from 30 per cent to 16 per cent; in timber from 13 per cent to 7 per cent; and, in iron ore from 12 per cent to 2 per cent. The loss of Africa's markets in cocoa beans, coffee and timber was largely to Asian countries; in iron ore to Latin American countries; and, in cotton to Eastern European countries.
21. See Hussain (1998).

22. See McCombie (1989 and 1992).
23. We used for the estimation only 9 commodity groups. However, if disaggregated data are used more stratification of elasticities will be possible.
24. See African Development Bank (2004).

REFERENCES

African Development Bank (ADB) (1995), *African Development Report 1995*, Abidjan, Côte d'Ivoire: ADB.
African Development Bank (ADB) (2003), *African Development Report 2003: Globalization and Africa's Development*, Oxford and New York: Oxford University Press.
African Development Bank (ADB) (2004), *African Development Report 2004: Africa in the Global Trading System*, Oxford and New York: Oxford University Press.
Andersen, P.S. (1993), 'The 45° Rule Revisited', *Applied Economics*, **25**(October), 1279–84.
Atesoglu, H.S. (1993), 'Balance of Payments Constrained Growth: Evidence from the United States', *Journal of Post Keynesian Economics*, **15**, 507–14.
Atesoglu, H.S. (1994a), 'Exports, Capital Inflows, Relative Price and Economic Growth in Canada', *Journal of Post Keynesian Economics*, **16**, 289–97.
Atesoglu, H.S. (1994b), 'Balance of Payments Determined Growth in Germany', *Applied Economics Letters*, **1**, 89–91.
Bairam, E. (1988), 'Balance of Payments, the Harrod Foreign Trade Multiplier and Economic Growth: the European and North American Experience, 1970–85', *Applied Economics*, **20**(12), December, 1635–42.
Bairam, E. and G.J. Dempster (1991), 'The Harrod Foreign Trade Multiplier and Economic Growth in Asian Countries', *Applied Economics*, **23**(11), November, 1719–24.
Chenery H. and I. Adelman (1966), 'Foreign Aid and Economic Development: The Case of Greece', *Review of Economic and Statistics*, **48**(February), 1–19.
Chenery H., I. Adelman and A. Strout (1966), 'Foreign Assistance and Economic Development.' *American Economic Review*, **56**(4), September, 679–733.
Dollar, D. and W. Easterly (1998), 'The Search for the Key: Aid, Investment and Policies in Africa', *Journal of African Economies*, **8**(4), December, 546–77.
Domar, E. (1947), 'Expansion and Employment', *American Economic Review*, **37**(1), March, 34–42.
Easterly, W. (1997), 'The Ghost of Financing Gap: How the Harrod–Domar Model Still Haunts Development Economics', *World Bank Policy Research Working Paper*, No. 1807 (July).
Economic Commission for Africa (ECA) (1999), *The Economic Report on Africa*, Addis Ababa: ECA.
Elliott, D.R. and R. Rhodd (1999), 'Explaining Growth Rate Differences in Highly Indebted Countries: An Extension to Thirlwall and Hussain', *Applied Economics*, **31**(9), September, 1145–48.
Harrod, R. (1939), 'An Essay in Economic in Dynamic Theory', *Economic Journal*, **49**(March), 14–33.
Hussain, M.N. (1997), 'Financial Liberalization, Currency Substitution and Investment: The Case of Egypt' *African Review of Money Finance and Banking*, **1–2**.
Hussain, M.N. (1998), 'Africa's External Sector and Economic Growth: Broad Pointers

for Development Assistance', in H. Kifle (ed.), *Towards a New Partnership for African Development*, Abidjan: African Development Bank.
Hussain, M.N. (1999), 'The Balance of Payment Constraint and Growth Rate Differences among African and East Asian Economies', *African Development Review*, **11**(1), June, 103–37.
Hussain, M.N. (2001), 'Exorcising the Ghost: An Alternative Growth Model for Measuring the Financing Gap', *Journal of Post Keynesian Economics*, **24**(1), Fall, 89–125.
Hussain, M.N. and A.P. Thirlwall.(1984), 'The IMF Supply-Side Approach To Devaluation: An Assessment with Reference to The Sudan', *Oxford Bulletin of Economics and Statistics*, **46**(May), 145–67.
Hussain, M.N. and A.P. Thirlwall (1986), 'The IMF Supply-Side Devaluation Approach: A Reply' *Oxford Bulletin of Economics and Statistics*, **48**(Feb), 83–6.
Hussain, M.N. and B.G. Gunter (2005) 'External Shocks and the HIPC Initiative: Impacts on Growth and Poverty in Africa' African Development Bank, *Economic Research Working Paper Series*, No. 75 (July).
Kaldor, N. (1975), 'Economic Growth and the Verdoorn Law – A Comment on Mr. Rowthorn's Article', *Economic Journal*, **85**(340), December, 891–6.
Kaldor, N. (1978), 'The Effects of Devaluation on Trade in Manufactures', in N. Kaldor, *Further Essays on Applied Economics*, London: Duckworth.
León-Ledesma, M.A. (1999), 'An Application of Thirlwall's Law to the Spanish Economy', *Journal of Post Keynesian Economics*, **21**(3), Spring, 431–9.
Lewis, W.A. (1954), 'Economic Development with Unlimited Supplies of Labour', *Manchester School of Economics and Social Studies*, **22**, 139–91.
Lucas, R. (1988), 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, **22**, 3–42.
McCombie, J.S.L. (1985), 'Economic Growth, the Harrod Foreign Trade Multiplier and the Hicks Super-multiplier', *Applied Economics*, **17**(1), February, 55–72.
McCombie, J.S.L. (1989), '"Thirlwall's Law" and Balance of Payments Constrained Growth – A Comment on the Debate', *Applied Economics*, **21**(5), May, 611–29.
McCombie, J.S.L. (1992), '"Thirwall's Law" and Balance of Payments Constrained Growth: More on the Debate', *Applied Economics*, **24**(May), 493–512.
McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance of Payments Constraint*, New York: St Martin's Press.
McCombie, J.S.L. and A.P. Thirlwall (1997), 'The Dynamic Harrod Foreign Trade Multiplier and the Demand-orientated Approach to Economic Growth: an Evaluation', *International Review of Applied Economics*, **11**(1) (January), 5–26.
McGregor, P.G. and J.K. Swales (1985), 'Professor Thirlwall and Balance of Payments Constrained Growth', *Applied Economics*, **17**(1), February, 17–32.
Mulligan, R.F. (1996), 'Export–Import Endogeneity in the Context of the Thirlwall–Hussain Model: an Application of the Durbin–Wu–Hausman test incorporating a Monte Carlo Experiment', *Applied Economic Letters*, **3**(4), March, 275–9.
Nashashibi, K. and P. Clawson (1986), 'The IMF Supply-Side Approach to Devaluation: A Response', *Oxford Bulletin of Economics and Statistics*, **48**(Feb), 73–82.
Perraton, J. (1990), 'The Harrod Foreign Trade Multiplier and the Developing Countries, 1970–1984: An Examination of the Thirlwall Hypothesis', University of Nottingham (memorandum).
Romer P.M. (1986), 'Increasing Returns and Long Run Growth', *Journal of Political Economy*, **94**(October), 1002–37.
Solow, R. (1956), 'A Contribution to the Theory of Economic Growth.' *Quarterly Journal of Economics*, **70**(February), 65–94.

Thirlwall, A.P. (1979), 'The Balance of Payments Constraint as an Explanation of International Growth Rate Differences', *Banca Nazionale del Lavoro Quarterly Review*, **128**, 44–53.
Thirlwall, A.P. (1986), 'Balance of Payments Constrained Growth: A Reply', *Applied Economics*, **18**(12), December, 1259–63.
Thirlwall, A.P. (1997), 'Reflections on the Concept of Balance-of-Payments Constrained Growth', *Journal of Post Keynesian Economics*, **19**(3), 377–85.
Thirlwall, A.P. and H.D. Gibson (1992), *Balance of Payments Theory and the United Kingdom Experience*, London: Macmillan.
Thirlwall, A.P. and N.M. Hussain (1982), 'The Balance-of-Payments Constraint, Capital Flows and Growth Rate Differences Between Developing Countries', *Oxford Economic Papers*, **34**(November), 498–510.
Williamson J. (1984), 'Is There an External Constraint?', *National Institute Economic Review*, **109**(August), 73–7.

2. Thirlwall's Law and Palley's pitfalls: a reconsideration

Mark Setterfield

1. INTRODUCTION

Thirlwall's Law, together with the idea that the potential rate of growth – Harrod's natural rate – is endogenous to the demand-determined actual rate of growth, rank amongst the most important of Tony Thirlwall's many contributions to the economics of growth and development. This chapter draws and expands on these contributions, paying particular attention to the 'pitfalls' in contemporary growth theory identified by Palley (2002). These pitfalls are associated with the common failure of both supply- and demand-led theories of growth to explicitly attend to the *reconciliation* of the actual and potential growth rates when employing steady state growth frameworks.

The remainder of the chapter is organized as follows. Section 2 briefly describes Thirlwall's Law and the balance-of-payments-constrained (BPC) growth model with which it is associated, whilst section 3 introduces 'Palley's pitfalls' and identifies their potentially damaging implications for the BPC growth model – even when the endogeneity of the natural rate of growth is taken into account. Section 4 then discusses different channels through which the actual and potential rates of growth might be reconciled in the BPC growth model. Palley's (2002) preferred solution is shown to involve quasi-supply-determined growth, and this is contrasted with an approach that results in a model of fully-demand-determined growth. Finally, section 5 offers some conclusions and suggestions for further research.

2. THIRLWALL'S LAW

The essence of Thirlwall's Law (Thirlwall, 1979) is that absent the ability to attract a permanent net inflow of capital from abroad, the rate of growth of an economy will be constrained by the requirement that it achieve current account balance, by selling sufficient exports to pay for its imports.[1] In its simplest

form, the BPC growth model that supports this insight can be stated as follows:[2]

$$x = a_0 g^* \quad [1]$$

$$m = b_0 g^a \quad [2]$$

$$g^a = g^d \quad [3]$$

$$x = m \quad [4]$$

where x is the rate of growth of exports, g^* is the rate of growth of foreign income, m is the rate of growth of imports, g^a is the actual rate of (domestic) growth, g^d is the rate of growth of aggregate demand (all measured in real terms) and a_0 and b_0 are the income elasticities of demand for exports and imports, respectively. Equations [1] and [2] are export and import growth equations,[3] equation [3] describes the actual rate of growth at any point in time as being demand-determined, and equation [4] imposes dynamic current account balance. Substituting [3] into [2] and the resulting expression, together with [1], into [4], we arrive at:

$$g^d = \frac{a_0 g^*}{b_0} \quad [5]$$

This is Thirlwall's Law, and constitutes the equilibrium BPC growth rate derived from equations [1]–[4]. Any effort to raise the rate of growth of demand (and hence the actual rate of growth) above the equilibrium value in [5] will result in $m > x$ and hence a current account deficit. The requirement that the current account should balance will thus force a reduction in g^d towards the equilibrium growth rate in [5] if growth is to be sustainable.[4]

The BPC growth model developed above can be extended in a variety of ways. As previously intimated, the rates of growth of relative prices and the exchange rate can be included amongst the determinants of x and m (McCombie and Thirlwall, 1994, pp. 234–7; McCombie and Roberts, 2002, pp. 90–91), whilst the balance of payments constraint in [4] can be modified to allow for permanent current account deficits financed by net inflows of foreign capital (McCombie and Thirlwall, 1994, pp. 246–9; McCombie and Roberts, 2002, pp. 93–6). The model can be extended to take account of the specific effects of interest payments to foreign creditors on the equilibrium BPC growth rate (Moreno-Brid, 2003), and can also be shown to generate chaotic dynamics rather than a unique and stable equilibrium outcome under certain circumstances (McCombie and Roberts, 2002, pp. 99–102).

There also exists a large empirical literature that tests and applies the BPC growth model, using data from both developed and less developed economies. McCombie and Thirlwall (1994, Chapter 3) and McCombie and Roberts (2002, pp. 96–8) survey this literature up to 1999; see Hussain (2001), Loria (2003), Moreno-Brid (2003) and Holland et al (2004) for examples of more recent tests and applications of the model. The obvious point to be made here is that in the twenty-five years since Thirlwall's (1979) seminal article, the BPC growth model has become a theoretically and empirically well established feature of the literature on growth and development.

3. PALLEY'S PITFALLS

i) The Problem

As is clear from the inclusion of equation [3] in the model developed above, Thirlwall's Law and the associated BPC growth model conceptualize growth as a demand-led process. This vision is, of course, at odds with the dominant neoclassical conceptualization of growth as a supply-led process.[5] But according to Palley (2002), both supply- and demand-led growth theories suffer a common problem: they neglect to explicitly discuss how the rates of growth of demand and supply (or alternatively, the actual and potential rates of growth) are reconciled.[6] Hence supply-led theories model factor accumulation and technical progress – i.e., the rate of growth of potential output – and then *implicitly* rely on a dynamic version of Say's Law to ensure that the rates of growth of demand and hence actual output converge towards the explicitly modelled rate of growth of supply. Demand-led theories, meanwhile, explicitly model the rates of growth of demand and hence actual output, and then *implicitly* rely on what Cornwall (1972, pp. 67–9) dubs 'Say's Law in reverse' to ensure that the potential rate of growth (i.e., the rate of growth of supply) converges towards the (explicitly modelled) rate of growth of demand. The problem with all this is that the reconciliation of the rates of growth of demand and supply is too important to be treated only by means of implicit theorizing. Hence absent the equality of the rates of growth of potential and actual output, the economy will experience either ever-increasing or ever-decreasing rates of capacity utilization. Apart from the problem that this would present for reconciling growth models with empirical evidence, it is, of course, a logical absurdity in models (including the BPC growth model) that rely on the concept of the steady state to describe growth outcomes, since the rate of capacity utilization is bounded above and below. The difficulty of this problem is compounded by the fact that efforts to explicitly consider how the rates of growth of supply and demand might be reconciled suggest that there may only be a discrete range of growth rates for which such

reconciliation is possible and for which the 'equilibrium' rate of growth is, therefore, sustainable as a steady state outcome (Cornwall, 1972; Palley, 1997).[7] Indeed, this range may be so small as to be confined to a single value of the growth rate.

ii) Implications for the BPC Growth Model

The idea that it is important to consider the *interaction* of the actual and potential rates of growth will not seem novel to anyone familiar with the work of Tony Thirlwall. As intimated earlier, one of Thirlwall's most important contributions to the economics of growth and development is his championing of the notion that the potential rate of growth, rather than constituting the exogenously given ceiling originally described by Harrod, much less the stable equilibrium growth rate envisaged by neoclassical growth theory, is actually *endogenous* to the (demand-determined) actual rate of growth (León-Ledesma and Thirlwall, 2000, 2002). But the idea that changes in the rate of growth of demand may induce changes in the rate of growth of supply does not suffice to avoid Palley's pitfalls: describing the *interaction* of the actual and potential rates of growth in this fashion does not address and solve the problem of their ultimate *reconciliation*.[8] This is illustrated by the following, extended BPC growth model:

$$x = a_0 g^* \qquad [1]$$

$$m = b_0 g^a \qquad [2]$$

$$g^a = g^d \qquad [3]$$

$$x = m \qquad [4]$$

$$\lambda = c_0 + c_1 g^a \qquad [6]$$

$$g^s = \lambda + n \qquad [7]$$

$$g^d = g^s \qquad [8]$$

where equations [1]–[4] are identical to those used in the previous section to illustrate Thirlwall's Law, λ is the rate of growth of labour productivity, n is the rate of growth of the labour force and g^s denotes the rate of growth of supply (potential output). Equations [6] and [7] explicitly introduce the supply side into our analysis, together with a Thirlwallian conception of the interaction between the rates of growth of supply and demand. This is done by describing the economy's potential rate of growth (equation [7]) as being endogenous to the

demand-determined actual rate of growth via the Verdoorn Law (equation [6]). Equation [8], meanwhile, imposes the constraint that the rates of growth of supply and demand must be equal in the steady state.

As demonstrated by Palley (2002, pp. 120–1), the model described above is over-determined. As we have already seen, solving equations [1]–[4] yields Thirlwall's Law:

$$g^d = \frac{a_0 g^*}{b_0} \qquad [5]$$

But solving equations [3] and [6]–[8] yields:

$$g^d = \frac{1}{1-c_1}(c_0 + n) = g^s \qquad [9]$$

Equation [9] describes the rate of growth of demand necessary to reconcile the actual and potential growth rates, *given* the interaction of these growth rates due to equation [6]. The rates of growth described by equations [5] and [9] will only be equal if:

$$g^* = \frac{b_0(c_0 + n)}{a_0(1-c_1)} \qquad [10]$$

This condition will only materialize by chance and is, therefore, unlikely to be observed. It is far more likely that we will observe either $g^* > b_0(c_0 + n)/a_0(1 - c_1)$, in which case the economy will experience a continually increasing rate of capacity utilization, or $g^* < b_0(c_0 + n)/a_0(1 - c_1)$, in which case the economy will experience a continually decreasing rate of capacity utilization. In either of these cases, for the reasons described earlier, the BPC growth rate in [5] cannot be sustained as a steady state equilibrium.

4. RECONCILING THE ACTUAL AND POTENTIAL RATES OF GROWTH IN THE BPC GROWTH MODEL

The problem identified above, while certainly important, does not strike a fatal blow to the BPC growth model. This is because multiple solutions to the problem can be identified, all of which are compatible with the basic structure of the model outlined in equations [1]–[4] and [6]–[8] above. The purpose of this section is to identify two such solutions and to discuss their relative merits in the context of the basic *vision* of the growth process encouraged by Thirlwall's Law and BPC growth theory.

i) Palley's Solution

Drawing on Palley (1996), Palley (2002, pp. 121–3) suggests a means of reconciling the actual and potential growth rates in the BPC growth model that involves positing the endogeneity of the income elasticity of demand for imports to the rate of capacity utilization. The basic idea – which hinges on the notion that a rising rate of capacity utilization will be associated with the increasing prevalence of bottlenecks in the domestic economy, causing the share of increases in income spent on imports to rise – can be illustrated by replacing equation [2] in the extended BPC growth model of the previous section with:

$$m = b_0(E)g^a \qquad [2a]$$

where $b_0' > 0$ and $E = Y_a/Y_p$ is a measure of capacity utilization based on the ratio of the levels of actual and potential real output (Y_a and Y_p respectively). Note that we must now rewrite [5] as:

$$g^d = \frac{a_0 g^*}{b_0(E)} \qquad [5a]$$

Now suppose that $g^* = g_1^* > b_0(E_1)(c_0 + n)/a_0(1 - c_1)$ initially, so that the equilibrium BPC growth rate derived from [5a] exceeds the rate of growth of demand consistent with the rate of growth of supply in [9] (i.e., $g^d > g^s$). This situation is illustrated in Figure 2.1 below. Now note that it follows from the definition of E (bearing in mind equation [3] and the notation used in equation [7] to describe the rate of growth of potential output) that:

$$\frac{\dot{E}}{E} = g^d - g^s$$

so that:

$$\dot{E} = E(g^d - g^s)$$

Hence whenever $g^d > g^s$, we have $\dot{E} > 0$. We will therefore observe $\dot{E} = E_1(g_1^d - g^s) > 0$ in response to the initial situation depicted in Figure 2.1, as a result of which the value of b_0 will rise, decreasing the slope of the $g^d = a_0 g^*/b_0(E)$ schedule. E and b_0 will continue to rise in this manner as long as the actual rate of growth exceeds the potential rate of growth or, in other words, until $E = E_2$ so that $g^d = a_0 g_1^*/b_0(E_2) = g_2^d$ in Figure 2.1. At this point, the actual and potential rates of growth are reconciled: $\dot{E} = E(g_2^d - g^s) = 0$ so that $\dot{b}_0 = 0$ and the economy achieves a (sustainable) steady state growth equilibrium given by:

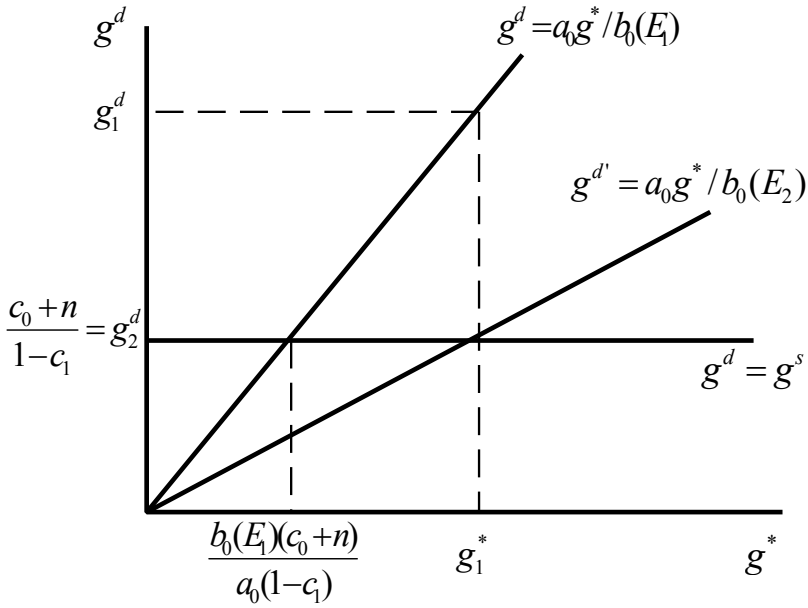


Figure 2.1 Reconciling the actual and potential rates of growth in the BPC growth model: Palley's solution

$$g_2^d = \frac{c_0 + n}{1 - c_1} = g^s < g_1^d$$

ii) An Alternative Solution

It will immediately be recognized that the steady state growth equilibrium above is identical to the rate of growth in equation [9], derived from the supply side of the model. In other words, the adjustment mechanism proposed by Palley (2002) for reconciling the actual and potential rates of growth in the BPC growth model involves changing the equilibrium BPC growth rate in equation [5a] (from g_1^d to g_2^d in the example above) 'to a level consistent with the underlying supply growth process' so that 'the steady-state growth rate … becomes uniquely determined on the supply-side' (Palley, 2002, p. 123). We have thus arrived at a model of *quasi-supply-determined growth*, in which the reconciliation of the actual and potential growth rates is achieved wholly by means of adjustments to the rate of growth of demand: the supply side 'rules the roost'.[9]

The question arises, then, as to whether there exist other processes capable of reconciling the actual and potential rates of growth in the BPC growth model that are, as it were, more sympathetic to the basic *vision* of this model, that ob-

served rates of growth are determined on the demand side. The answer is that there are. One such process involves treating the Verdoorn coefficient (c_1 in equation [6]) as endogenous to the rate of capacity utilization.[10] The Verdoorn coefficient captures the extent to which productivity growth is sensitive to the actual rate of growth of output, as a result of static and dynamic increasing returns. The hypothesis here is that the extent to which any given rate of output growth will induce productivity growth (i.e., the precise size of the Verdoorn coefficient) is a direct function of the rate of capacity utilization: firms will less likely be induced to engage in the innovation, technical change and organizational change from which productivity gains materialize by any given *rate of growth* of demand if the *level* of demand is low relative to that required for full capacity utilization. In other words, more productivity growth is induced by a goods market that is *both* tight *and* rapidly expanding. Empirical evidence for this claim can be found in the work of Cornwall and Cornwall (2001, Chapter 10), who show that both the rate of growth of demand and the extent of 'economic slack' (the disparity between actual and potential output and employment) are important determinants of productivity growth in the OECD.

Incorporating this insight into the BPC growth model developed earlier involves replacing equation [6] with:

$$\lambda = c_0 + c_1(E)g^a \qquad [6a]$$

where $c_1' > 0$ and E is as previously defined. Note that we must now rewrite [9] as:

$$g^d = \frac{c_0 + n}{1 - c_1(E)} = g^s \qquad [9a]$$

Now suppose that $g^* = g_1^* > b_0(c_0 + n)/a_0(1 - c_1(E_1))$ initially – i.e., once again we assume that the equilibrium BPC growth rate (this time derived from equation [5]) exceeds the rate of growth of demand consistent with the rate of growth of supply (that is now given by equation [9a] with $E = E_1$), so that $g_1^d > g_1^s = (c_0 + n)/(1 - c_1(E_1))$. This situation is illustrated in Figure 2.2 below. As before, we will observe $\dot{E} = E_1(g_1^d - g_1^s) > 0$ in this initial situation. But as a result of increasing rates of capacity utilization, it is now the value of c_1 that will rise: firms will become more willing to engage in innovation, technical change and organizational change as the goods market tightens, raising the rate of productivity growth associated with any given rate of output growth. This, in turn, will raise the rate of growth of potential output, causing the $g^d = g^s$ schedule in Figure 2.2 to shift upwards. E and c_1 will continue to rise as a long as $g_1^d > g^s$ or, in other words, until $E = E_2$ so that $g^s = g_2^s = (c_0 + n)/(1 - c_1(E_2))$. At this point, the actual and potential rates of growth are reconciled: $\dot{E} = E(g_1^d - g_2^s) = 0$ so that $\dot{c}_1 = 0$

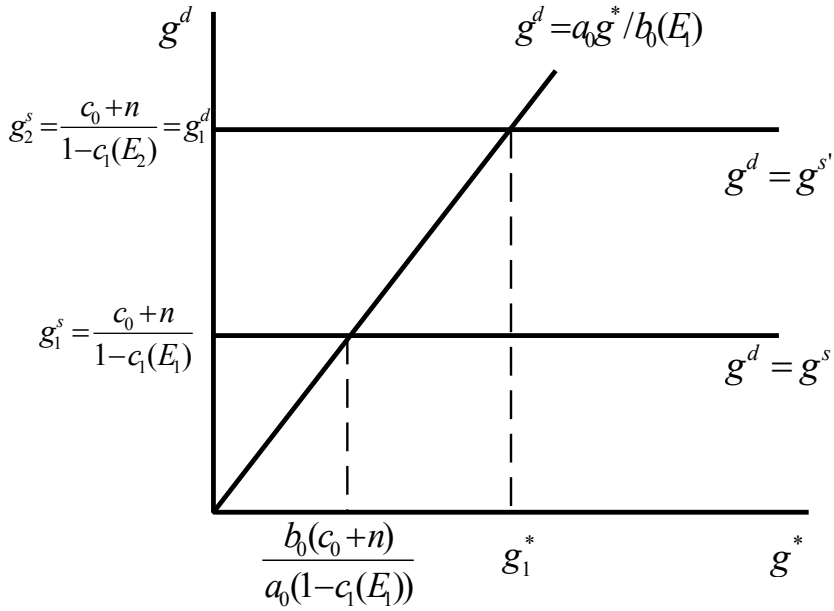


Figure 2.2 Reconciling the actual and potential rates of growth in the BPC growth model: an alternative solution

and the economy once again achieves a (sustainable) steady state growth equilibrium, this time given by:

$$g_1^d = \frac{c_0 + n}{1 - c_1(E_2)} = g_2^s > g_1^s$$

Inspection of the steady state growth equilibrium above reveals that the adjustment process that reconciles the actual and potential rates of growth now involves variations in the potential rate of growth to accommodate the unchanging (given the value of g^*) equilibrium BPC growth rate derived from equation [5]. In other words, it is the structure of the supply side that adjusts to accommodate the demand-determined actual rate of growth. The demand-side thus 'rules the roost' in what can be identified as a model of *fully-demand-determined growth*.[11]

iii) Some Further Remarks

At this point, two further remarks on the adjustment processes discussed above are warranted. First, note that despite the substantive difference in outcomes to

which the adjustment processes give rise (quasi-supply-determined versus fully-demand-determined growth rates), these adjustment processes are not mutually exclusive. It is entirely possible that *both* processes (and possibly even others beside – on which, again, see Palley (2002, pp. 123–4)) – will be operative at the same time. In this case, of course, the question of whether the final steady state rate of growth that emerges from the reconciliation of the actual and potential growth rates will be closer to the initial equilibrium BPC growth rate in [5a] or the initial rate of growth derived from the structure of the supply side in [9a] depends on the relative magnitudes of b_0' and c_1'.

Second, it will not have gone unnoticed that *both* adjustment processes are premised on a common conception of the growth process that involves the long run endogeneity of the rate of capacity utilization. The question of whether or not the rate of capacity utilization can be regarded as endogenous in the long run is, of course, a controversial topic in macrodynamics.[12] But even leaving this debate aside, the role of the rate of capacity utilization in both of the adjustment processes described above is significant because, as remarked earlier, this variable is bounded above and below. There are, therefore, strict *bounds* within which either adjustment process can be relied upon to reconcile the actual and potential rates of growth. If this reconciliation is not achieved before $E = 0$ or $E = 1$ (or, more likely, before the value of E reaches some economically – rather than logically – defined minimum/maximum value), then the problems identified in section 3(i) associated with the unsustainability of the equilibrium rate of growth will re-emerge.

5. CONCLUSION

The purpose of this chapter has been to further advance the development of the BPC growth model and discussion of the interaction between the actual and potential rates of growth – two of Tony Thirlwall's most important contributions to the economics of growth and development – by explicitly describing processes through which the actual and potential rates of growth can be *reconciled* in the BPC growth model. Concern with this issue is not new, having previously been voiced by Palley (2002). But the process that Palley proposes for reconciling the actual and potential rates of growth in a BPC growth framework gives rise to a model of quasi-supply-determined growth. The alternative process proposed in this chapter places the burden of adjustment on the supply side, making the equilibrium BPC growth rate the ultimate determinant of the steady state growth rate within a model of fully-demand-determined growth.

As has previously been remarked, the adjustment processes proposed by Palley (2002) and in section 4(ii) above are not mutually exclusive. Moreover, both are dependent on the long run endogeneity of the rate of capacity utilization

– a controversial issue in macrodynamics, and one that draws attention to the existence of strict bounds on the efficacy of either adjustment process. Obviously, the task of further exploring which of the adjustment processes discussed above is most important, and whether there is sufficient variability in the rate of capacity utilization in the long run to accommodate the workings of *either* process, must fall to empirical investigation. Such analysis might, therefore, be identified as the next step for those seeking to add to the already considerable body of evidence supporting the relevance of Thirlwall's Law and the associated BPC growth model as explanations of long run growth.

NOTES

1. The model can be extended to allow for permanent trade deficits financed by capital inflows, in which case the ultimate constraint becomes the inability of an economy to continuously increase its foreign debt to GDP ratio. See, for example, McCombie and Thirlwall (1994, pp. 246–9) and McCombie and Roberts (2002, pp. 93–6). This extension is overlooked in what follows for the sake of simplicity.
2. The notation used below follows that of Palley (2002) in order to better facilitate comparison of what follows with Palley's original discussion of the problem identified in section 3 and its solution in the context of the BPC growth model.
3. It is conventional to include relative prices and the exchange rate as arguments in the export and import demand functions from which [1] and [2] are implicitly derived, but these factors are again overlooked for the sake of simplicity. See also McCombie and Roberts (2002, pp. 90–1) and especially McCombie and Thirlwall (1994) on the notion that relative prices and the exchange rate are less important determinants of long run growth than the variables that enter into the determination of Thirlwall's Law.
4. By the same token, efforts to accumulate current account surpluses by growing at a slower rate than that indicated in [4] will impose deficits on the rest of the world, as a result of which we would expect g^* to fall (as foreign nations seek to restore current account balance) until the condition in [4] is once again satisfied. See Thirlwall (1997, 2001) and McCombie and Thirlwall (1994) for discussion of the links between Thirlwall's Law and BPC growth theory on one hand, and the contributions of Harrod and Kaldor to the economics of trade and growth on the other.
5. Both first generation (the Solow model) and second generation (models of endogenous growth) neoclassical growth theories share this preoccupation with the supply side.
6. Palley (2002) is, in fact, a reprint of a paper that was published in the *Review of Political Economy*, vol. 15, no. 1 (January), 75–84 (2003). All page references above are to the reprinted article.
7. See also Palley, 2002, pp. 117–19 for a brief summary of these results.
8. Indeed, according to Palley (2002, p. 115), it serves only to draw attention to the problem.
9. The Palley model is described above as involving *quasi*-supply-determined growth because the supply side is not autonomous in this model (as in neoclassical growth theory): the (demand-determined) actual rate of growth influences the potential rate of growth via the Verdoorn Law (equation [6]). Note also that, as Palley (2002, p. 123) himself points out, the steady state rate of *capacity utilization* in his model is demand-determined. Hence note that the steady state growth rate derived in the previous section implies that:

$$\frac{a_0 g_1^*}{b_0(E_2)} = \frac{c_0 + n}{1 - c_1}$$

so that:

$$E_2 = b_0^{-1}\left(\frac{a_0 g_1^*(1-c_1)}{c_0+n}\right)$$

The rate of growth of demand for exports ($x_1 = a_0 g_1^*$) is thus a determinant of the steady state rate of capacity utilization.

10. It should be noted that Palley (2002, pp. 123–4) himself briefly draws attention to a second process, operating through the endogeneity of the rate of growth of the labour force.
11. Note that, as in Palley's model of quasi-supply-determined growth, the steady state rate of capacity utilization is again demand-determined. This can be verified by noting that the steady state growth equilibrium derived above implies that:

$$\frac{a_0 g_1^*}{b_0} = \frac{c_0+n}{1-c_1(E_2)}$$

So that:

$$E_2 = c_1^{-1}\left(1-\frac{b_0(c_0+n)}{a_0 g_1^*}\right)$$

Once again, the rate of growth of demand for exports ($x_1 = a_0 g_1^*$) is revealed to be a determinant of the steady state rate of capacity utilization.

12. See, for example, Lavoie (1995).

REFERENCES

Cornwall, J. (1972), *Growth and Stability in a Mature Economy*, London: Martin Robertson.

Cornwall, J. and W. Cornwall (2001), *Capitalist Development in the Twentieth Century: an Evolutionary-Keynesian Analysis*, Cambridge: Cambridge University Press.

Holland, M., F.V. Vieira and O. Canuto (2004), 'Economic growth and the balance-of-payments constraint in Latin America', *Investigacion Economica*, **63**, 45–74.

Hussain, M.N. (2001), '"Exorcising the ghost": an alternative model for measuring the financing gap in developing countries', *Journal of Post Keynesian Economics*, **24**, 89–124.

Lavoie, M. (1995), 'The Kaleckian model of growth and distribution and its Neo-Ricardian and Neo-Marxian critiques', *Cambridge Journal of Economics*, **19**, 789–818.

León-Ledesma, M. and A.P. Thirlwall (2000), 'Is the natural rate of growth exogenous?', *Banca Nazionale Del Lavoro Quarterly Review*, **215**, 433–45.

León-Ledesma, M. and A.P. Thirlwall (2002), 'The endogeneity of the natural rate of growth', *Cambridge Journal of Economics*, **26**, 441–59.

Loria, E. (2003), 'The Mexican economy: balance-of-payments constrained growth model – the importance of the exchange rate, 1970–1999', *Journal of Post Keynesian Economics*, **25**, 661–91.

McCombie, J.S.L. and M. Roberts (2002), 'The role of the balance of payments in economic growth', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance-of-Payments Constraint*, London: Macmillan.
Moreno-Brid, J.C. (2003), 'Capital flows, interest payments and the balance-of-payments constrained growth model: a theoretical and empirical analysis', *Metroeconomica*, **54**, 346–65.
Palley, T.I. (1996), 'Growth theory in a Keynesian mode', *Journal of Post Keynesian Economics*, **19**, 113–35.
Palley, T.I. (1997), 'Aggregate demand and endogenous growth: a generalized Keynes–Kaldor model of economic growth', *Metroeconomica*, **48**, 161–76.
Palley, T.I. (2002), 'Pitfalls in the theory of growth: an application to the balance-of-paymentsconstrained growth model', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (1979), 'The balance of payments constraint as an explanation of international growth rate differences', *Banca Nazionale Del Lavoro Quarterly Review*, **128**, 45–53.
Thirlwall, A.P. (1997), 'Reflections on the concept of balance-of-payments-constrained growth', *Journal of Post Keynesian Economics*, **19**, 375–85.
Thirlwall, A.P. (2001), 'The relation between the warranted rate of growth, the natural rate and the balance-of-payments equilibrium growth rate', *Journal of Post Keynesian Economics*, **24**, 81–8.

3. On specifying the demand for imports in macroeconomic models

G.C. Harcourt

When I wrote a book note for the *Economic Journal* in 1998 on Volume 2 of Tony Thirlwall's splendid *Selected Essays, Macroeconomic Issues From a Keynesian Perspective* (1997), I said:

> With Tony Thirlwall what you see is what you get. He has strong views, well thought out and stuck to, he is lucid, humane and persuasive to those who have eyes to see and ears to hear – not many of us left, unfortunately. He is proud to be called, and to call himself, 'an unreconstructed Keynesian'. He is; he is also an original and innovative economist who, within a broadly Keynesian framework, has with his papers illuminated our understanding of some of the most pressing modern issues.

Of course, the content, scope and depth of Tony's essays should have warranted a review, preferably a review article, but such is the technocratic and Philistine character of our age and profession that an appraisal of his contributions in the public domain is almost invisible, confined, as it is, to a short book note only available on the web. (Book notes are no longer published in the journal itself.)

In one of his essays on 'the input-output formulation of the foreign trade multiplier', Tony points out that together with himself only Wonnacott (1974) and Harcourt, Karmel and Wallace (1967) treat the demand for imports in macroeconomic analysis as a function of Gross Expenditure (that is to say, gross of expenditure on imports) rather than of National Income. Even then, Wonnacott and Harcourt, Karmel and Wallace barely scratch the surface of such a treatment. By contrast, in Tony's essay (written with the late Charles Kennedy), their treatment is allied with an analysis of the movements of imports through the input-output structure of the economy.

These considerations prompted this essay for Tony's volume, to which I feel privileged and proud to contribute. I want to investigate some more the *raison d'être* for treating imports as a function of expenditure rather than income and also refer to the links between Keynesian macroeconomic national accounting concepts and categories, and the production interdependent models of Leontief and Sraffa.

I first wrote on the second topic when I was working on the first draft of a chapter on the national accounts in *Economic Activity* (Chapter 2). At much the same time I was also working with Vincent Massaro on Piero Sraffa's 1960 classic, *Production of Commodities by Means of Commodities*, especially on Sraffa's ingenious device of sub-systems, see Harcourt and Massaro (1964). It was through these two (non-trivial) pursuits that I obtained at least a glimpse of the nature of the complex relationships between the respective systems of Keynes and Sraffa, so deftly and deeply developed by Tony and Charles Kennedy in their essay. But first I want to investigate the implications of making imports a function of expenditure and try to develop a simple quantitative estimate of the bias introduced into macroeconomic models by making imports a function of income. I then question both procedures in the light of Keynes's two concepts of the aggregate demand function in *The General Theory* – see pp. 24, 25, 29, 30. I hope these musings, albeit incomplete and inconclusive, will nevertheless be of interest to Tony, who has that indispensable characteristic of the really great economist, to wit, to know when perhaps definite answers may not be given to important questions in economics, a lesson I learnt very early on from my greatest Australian mentor, Eric Russell. Incidentally, it was Eric who led the way on the Australian scene in making imports a function of expenditure.

Why is it the most common practice to make the demand for imports a function of income rather than expenditure? I suppose the most convincing rationale must be that imports, like saving and taxation, are all leakages from the flows of expenditure, production and income that need to be offset by injections of investment, government and export expenditures in order to establish the short-period equilibrium flow of national income, production and expenditure. Moreover, saving and taxation are self-evidently leakages from the received income of the community. Therefore, in modelling an open economy with a government sector and deriving an expression for the full multiplier at work, symmetry requires that the three marginal propensities reflecting the three leakages be added together in the multiplier formulation: $1/(s + t + m)$. However, while this reasoning seems valid for saving and taxation, is it obviously so for imports as well? No, because imports are demanded in any short period, when techniques of production, relative prices and exchange rates may be taken as given, in order to make production possible in response to anticipated (or even known) sales. Therefore, the level of final expenditure on commodities associated with consumption, investment, government, and export categories of expenditure seems the obvious place to go to in order to find what sectoral and total demand for imports are likely to be in any given short-period situation. Of course, the bulk of activity takes place in firms that produce intermediate goods, not final goods, and so we must conceive of demands being passed down the line, as it were, in the interdependent production processes that characterise the

structure of modern economies. This is implied in the simple Keynesian constructions of the models in, say, *Economic Activity* where total imports are made a simple proportional function of total gross expenditure. As a simplification, no distinction is made between the import requirements associated with the production of the different categories of commodities. It is acknowledged, of course, that there may well be more imports required per unit of production, of, say, consumption, investment or even export commodities than those associated with much government expenditure, for example. This is regarded as a complication to be brought in later once the essential principles have been understood and the analysis worked through. Furthermore, by making imports a function of expenditure, the import requirements in production needed to match autonomous expenditures are explicitly taken into account in the formula for the open economy multiplier.

Suppose then that we accept that expenditure is the principal determinant of import requirements, that it signals what imports will be needed in any given short-period situation to produce the commodities to match expenditure. Does it make any essential difference to the ultimate answers that we give for the determination of levels of production and income whether we have used expenditure or income as the determinants of import demand?

The answer is that at *short-period equilibrium* there is no difference. At equilibrium aggregate demand equals aggregate supply (and total planned injections are equal to total induced leakages), so we must get the same overall answer. (We have to make allowance for the fact that the marginal propensity to spend on imports from expenditure has a different numerical value to the marginal propensity to import from income, because income by definition is always less than expenditure; but one is reducible to equivalence with the other in order to determine the same overall demand for imports.) This is clear in Figure 3.1, which is based on the diagrams in Chapter 12 of *Economic Activity*.

On the vertical axis we measure expenditure (E) and income (Y), on the horizontal axis, Y, all in real terms. We take $\overline{I}$, $\overline{G}$, and $\overline{X}$ to be autonomous and C to be a function of Personal Disposable Income (Y_d), $C = C(Y_d)$, with changes in C induced by changes in Y_d. (There is an autonomous term in the consumption function and in the short period a unique relationship between Y_d and Y so that C may be related in a derived way to Y.) m is the marginal propensity to spend on imports and so *domestic* expenditure is $(1 - m)(E)$, where $E = C(Y_d) + \overline{I} + \overline{G} + \overline{X}$.

The difference between the two lines is the expenditure on imports at each level of income. Not only is the 45° line a reference point, it may also be interpreted as an aggregate supply (and demand) line, plotting expected demand for *domestic* production against domestic production itself. (All the microeconomic activities in different market structures go on behind the line, as it were, as individual businesspeople react to expected prices, as in Keynes's *General Theory*,

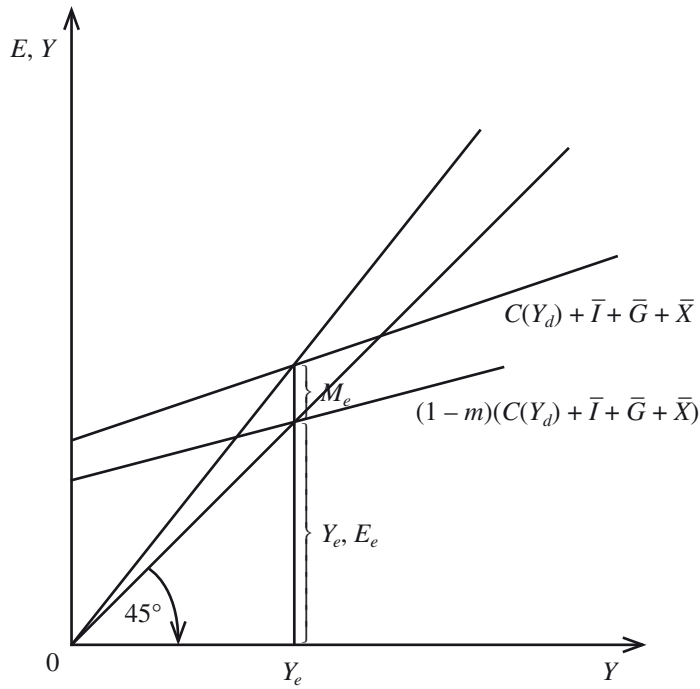


Figure 3.1 Short-period equilibrium in an open economy with a government

or to expected sales, when we have more realistic imperfectly competitive, monopolistic competitive and oligopolistic structures, and decide on short-period production and employment, creating the incomes that lead to consumption spending, saving and tax payments.) Then for each level of domestic demand and production to be possible in the short period, there must be accompanying inputs of imports. This is shown by the straight line starting at the origin and situated *above* the 45° line by the amount of imports needed, determined by the marginal propensity to import associated with each level of, and change in, Y. At the equilibrium level of Y, OY_e, which is determined by the intersection of $(1 - m)(E)$ with the 45° line, *both* methods of estimating the demand for imports give the same answer, *QED*!

Away from the equilibrium level, though, the answers differ. The demand for imports is estimated to be greater by the expenditure method (M_E) than by the income method (M_Y) to the *left* of the equilibrium intersection points, and to be less, to the right of the intersection, see Figure 3.2.

Is it possible to say anything more quantitative about this discrepancy? We know that in a very simple model, if we suppose prices to be given and constant

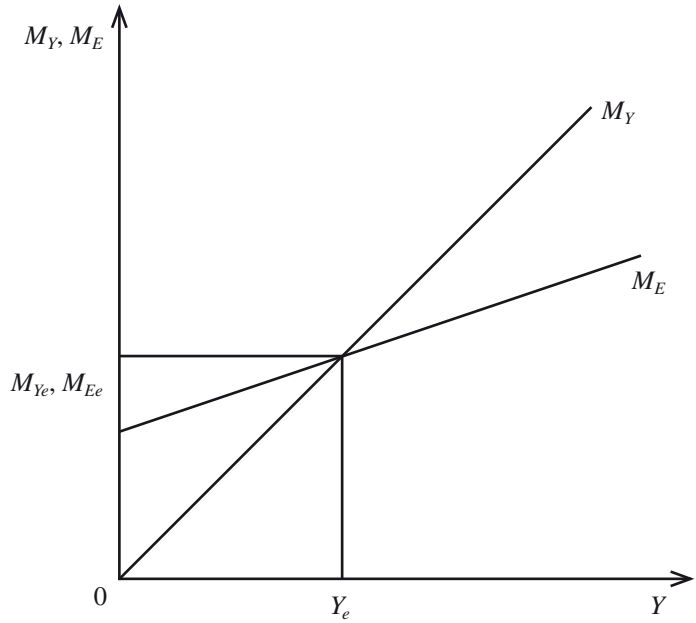


Figure 3.2 Demands for imports by the expenditure and income methods

in the short period, then to the left of the intersection of the aggregate demand function with the aggregate supply function (as we have defined them), we have an unintended run-down of stocks (or a failure to add to them as much as planned) because aggregate demand exceeds aggregate supply (planned investment exceeds planned saving); to the right of the intersection, we have an unintended build-up of stocks because of the excess supply situations which characterise these positions. Then, a moment's reflection suggests (when I told Joan Robinson this she said it was a *very* subtle point) that the difference in the demand for imports predicted by the two methods is equal to *the import components of the unintended changes in stocks occurring*. To the left of the equilibrium point, if we accept for the moment that the expenditure method is the correct way to specify the demand for imports, the income method results in an *underestimate* of the total demand for imports by these amounts; to the right of the equilibrium point, it results in an overestimate. So perhaps this simple result could be taken as a starting point for actual estimates of total imports in given periods of actual time, if it is accepted that the economy is rarely if ever to be found at its short-period equilibrium level. Elsewhere (see Harcourt 1981; 1982), I have argued that the short-period equilibrium level is a sort of centre of gravitation, a sometimes justified short cut for viewing actual national

income and expenditure figures as though they were identical with the levels corresponding to their centres of gravitation; in which case no bias would need to be estimated.

Until I started to think about this essay for Tony's volume, I must admit I thought that where we have now reached was the end of the matter. Now I am not so sure. The reason why is as follows: as we mentioned earlier, there are two different concepts of the aggregate demand function in *The General Theory*, only one of which has on the whole survived in later writings by Keynes himself and others, and in the textbooks. The first, which I think is more fundamental, is the summation of what businesspeople think the sales of their products will be in any given short period. It is these expectations which guide their production planning, employment offers and, I now want to argue, *their orders for imports*. Some businesspeople will produce only to order and will know exactly what employment to offer and imports to buy. Others, whether they sell final goods or supply intermediate goods, have to anticipate what their sales will be and therefore what production to plan, employment to offer and imports to buy. Keynes put it like this (he ignored imports, of course).

> … in a given situation of technique, resources and factor cost per unit of employment, the amount of employment, both in each individual firm and industry, and in the aggregate, depends on the amount of proceeds which the entrepreneurs expect to receive from the corresponding output … let D be the proceeds which entrepreneurs expects to receive from the employment of N men, the relationship between D and N being written $D = f(N)$, which is called the *aggregate demand function*. (1936; *C.W.*, vol VII, 1973, 24-25, emphasis in original)

If we suppose that included in these decisions are sales to stock in their own businesses, then the total of *all* these expectations of sales by business people must be *exactly* equal to the production being planned and implemented and therefore the total income that is created (it would be irrational for them to produce either more or less). Hence the 45° line is both a reference point *and* the aggregate demand (and supply) function. It is the aggregate supply function because, remembering our assumptions, it corresponds with Keynes's definition:

> … the aggregate supply price of the output of a given amount of employment is the expectation of proceeds which would just make it worth the while of the entrepreneurs to give that employment … let Z be the aggregate supply price of the output from employing N men, the relationship between Z and N being written $Z = \phi(N)$, which can be called the *aggregate supply function*'. (1936; *C.W.*, vol. VII, 1973, 24-25, emphasis in original)

Of course, whether the individual and total expectations are realised or not depends upon the subsequent reactions of the income receivers with regard to

their spending on consumption goods as indicated by the aggregate consumption function, which is a major component of the second concept of the aggregate demand function in *The General Theory*, that is to say, the sum of planned expenditures on all final goods, either related to income levels or autonomously in the simple case here for $\overline{I}$, $\overline{G}$ and $\overline{X}$. This interpretation is implied in Keynes's summing up of his theory on pp. 28–32, especially in (3) and (5), p. 29:

> (3) The amount of labour N which the entrepreneurs decide to employ depends on the sum (D) of *two* quantities, namely D_1, the amount which the community is expected to spend on consumption, and D_2, the amount which it is expected to devote to new investment …
> […]
> (5) Hence the volume of employment in equilibrium depends on (i) the aggregate supply function, ϕ, (ii) the propensity to consume, χ, and (iii) the volume of investment, D_2 … the essence of the General Theory of Employment'. (emphasis in original)

The latter form of the aggregate demand function does not in principle have to be known by individual businesspeople who, of course, are mainly interested in what their specific prices and/or sales will be. But it is known to us, the all-seeing macroeconomists, looking at the economy as a whole. And it is for this reason, as well as others, that import demands were related to expenditure in the models of the three sets of authors mentioned above.

But were they correct to do so? Probably not, because if the argument above concerning the first concept of the aggregate demand function is correct, then the *initiation* of production, employment and demand for imports must come from the combined (but uncoordinated) actions of all the businesspeople in the community. In which case, the demand for imports will be related to points on the 45° line, not to points on the expenditure relationship which is derived from the second concept (onto which Keynes quickly moved). So the more common method of estimating the demand for imports by making it a function of income will give the right answer, though up until now, says he modestly, exposing former treason, for the wrong reason.[1] Moreover, there will be no reason to correct for biases in the estimates of the demand for imports, as was suggested above, because income, *whether equilibrium or not*, is the relevant major determinant of such demand in the short period.

I shall be the first to cheer if Tony shows that I am now wrong and that he and Kennedy, in a much more detailed and persuasive way, and the other authors were right all the time. So, over to you, Tony, and many congratulations on your splendid contributions as a wise, unreconstructed Keynesian!

NOTE

1. 'The last temptation is the greatest treason:
   To do the right deed for the wrong reason.'
   Eliot (1935), 44.

REFERENCES

Eliot, T.S. (1935), *Murder in the Cathedral*, London: Faber and Faber.
Harcourt, G.C. (1981), 'Marshall, Sraffa and Keynes: Incompatible Bedfellows?', *Eastern Economic Journal*, **5**, 39–50. Reprinted in Harcourt (1982).
Harcourt, G.C. (1982), *The Social Science Imperialists. Selected Essays*, ed. by Prue Kerr, London: Routledge and Kegan Paul. Reprinted in the Routledge Library Editions series in 2003.
Harcourt, G.C., P.H. Karmel and R.H. Wallace (1967), *Economic Activity*, Cambridge: Cambridge University Press.
Harcourt, G.C. and Vincent G. Massaro (1964), 'A Note on Mr Sraffa's Sub-Systems', *Economic Journal*, **74**, 715–22.
Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London: Macmillan, *C.W.*, vol. VII, 1973.
Sraffa, P. (1960), *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*, Cambridge: Cambridge University Press.
Thirlwall A.P. (1997), *Macroeconomic Issues from a Keynesian Perspective. Selected Essays of A.P. Thirlwall, Volume Two*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Wonnacott, P. (1975), *Macroeconomics*, London: Irwin.

4. Keynes, Post Keynesian analysis, and the open economies of the twenty-first century

Paul Davidson

INTRODUCTION

It is often said by mainstream economists that Post Keynesian theory's contribution to economic theory is all negative in that it attacks orthodox (supply side) theory without providing a positive alternative coherent contribution. Moreover, it is claimed that this failure is due to the fact that Post Keynesian theory is based on the closed economy analysis of Keynes's *General Theory*, while the real world economies of the 21st century are open to world trade. Keynes might have had some relevance for the developed economies of the mid-20th century that were enmeshed in the Great Depression, but Keynes has little or nothing to offer to the open economies of a globalized economic system of the 21st century.

These disparaging statements regarding both Keynes and the Post Keynesian analytical framework are wholly false and merely represent the best barbs that mainstream classical economists, who have been stung by the rapier of Post Keynesian analysis, can muster.

Section I of this paper will cite chapter and verse in Keynes's *General Theory* where insights and policies are suggested that are relevant today to nations in a globalized economic world. Section II briefly indicates how the orthodox theory of the efficiency of flexible exchange rates fails to meet the test of the facts. Section III focuses on one enormous positive contribution made by Tony Thirlwall that permits the analyst to understand the income effects of world trade, including the potential widening income inequalities between developed nations and those less developed countries that pursue economic industrial policies based on the law of comparative advantage. Tony Thirlwall deserves kudos for developing this analysis from the Keynes-Harrod multiplier literature of the 1930s. When mainstream economists finally recognize the importance of income effects of trade on open economies, then Tony Thirlwall's name will be cited for developing the pioneering work that Post Keynesians call Thirlwall's Law.

KEYNES AND THE OPEN ECONOMY

It is a paradox that Keynes, a man who spent most of his professional life working on the analysis of open economies, should be thought of as having developed a general theory that is only applicable to closed economies. It is true that much of Keynes's *General Theory of Employment Interest and Money*[1] is developed in a closed economy context. One possible reason for this closed economy focus is that Keynes wanted to demonstrate that even if one abstracted from international trade complications, a closed economy did not possess any automatic market mechanism that assured a full-employment equilibrium. Once this was demonstrated, however, Keynes did introduce some major open economy aspects in his general theory when he noted that

a. Trade could modify the magnitude of the domestic employment multiplier [Keynes, 1936, p. 120],
b. reductions in money wages would worsen the terms of trade and therefore reduce real income, though it could improve the balance of trade payments [Keynes, 1936, p. 263], and
c. stimulating either domestic investment or foreign investment could increase domestic employment growth [Keynes, 1936, p. 335].

In our 21st century economic system, under the influence of a resurgent classical economic analysis by mainstream economists, governments are afraid that deliberately stimulating any domestic component of aggregate demand via budget deficits will unleash inflationary forces. Accordingly, export-led growth is seen as the only alternative path for expanding domestic employment. A 'favorable balance [of trade], provided it is not too large, will prove extremely stimulating' to domestic employment [Keynes, 1936, p. 338], even if it does so at the expense of employment opportunities abroad.

In a passage that is particularly *a propos* to today's global economic setting, Keynes [1936 p. 335] noted that 'in a society where there is no question of direct investment under the aegis of public authority [due to fear of government deficits *per se*], the economic objects, with which it is reasonable for the government to be preoccupied, are the domestic interest rate and the balance of foreign trade'. If, however, nations permit free movement of capital funds across national boundaries, then 'the authorities had no direct control over the domestic rate of interest or the other inducements to home investment, [and] measures to increase the favorable balance of trade [are]…the only *direct* means at their disposal for increasing foreign investment' [Keynes, 1936, p. 336] and domestic employment.

Keynes explicitly noted that the domestic employment advantage gained by any government pursuing an export-led growth policy 'is liable to involve an

equal disadvantage to some other country' [Keynes, 1936, p. 338]. When many countries simultaneously pursue an 'immoderate policy' [Keynes, 1936, p. 338] of export-led growth (e.g., Japan, Germany and the NICs of Asia in the 1980s and Japan and China today), this aggravates the unemployment problem for these nations' trading partners.[2] Consequently, the trading partners of nations deliberately pursuing export-led growth policies are forced to engage in a 'senseless international competition for a favorable balance which injures all alike' [Keynes, 1936, pp. 338–9].

Any deliberate policy that aims to make a nation's industries more competitive in order to improve the trade balance requires either (a) forcing down nominal wages (including fringe benefits) to reduce labor production costs and/or (b) a devaluation of its exchange rate[3] (or an appreciation of its trading partners's exchange rate). Competitive gains obtained by manipulating these wage or exchange rate variables can only foster further global stagnation and recession as one nation's attempt to regain a competitive edge via (a) or (b) passes the black queen of reduced profits and higher unemployment to other nations in the global economy.

Unlike the classical theorists of his day (and our day as well[4]) Keynes recognized that 'the mercantilists were aware of the fallacy of cheapness and the danger that excessive competition may turn the terms of trade against a country' [Keynes, 1936, p. 345] thereby reducing domestic living standards. In an open economy, turning the terms of trade against a country can mean that domestic workers may be working more while earning less in real terms.

Keynes realized that in an open global economy the only path to global full employment might require *every* nation to actively and independently undertake a program for public *domestic* investment to generate *domestic* full employment. Otherwise, the resulting laissez-faire system of 'prudent' fiscal finance in tandem with a system of free international monetary flows would create a global environment where *each* nation could solve its unemployment problem only by seeking international competitive advantages resulting in export-led growth. The pursuit of such policies simultaneously by many nations, however, 'injures all alike' [Keynes, 1936, pp. 338–9].

This warning from Keynes, however, went virtually unrecognized in the last decades of the 20th century when mainstream economists waxed enthusiastically about the export-led economic miracles of Japan, Germany and the Pacific rim Newly Industrialized Countries – and later China and India – without noting that these miraculous performances were at the expense of the rest of the world.

Whenever governments do not have the political will to stimulate directly any domestic component of aggregate spending to reduce unemployment, 'domestic prosperity [is] directly dependent on a competitive pursuit of [export] markets' [Keynes, 1936, p. 349]. This is a competition in which all nations cannot be winners, but all *can* become losers!

For each nation simultaneously to break out of a global slow-growth stagnating economic environment and promote global prosperity, the correct policy, Keynes insisted, is not to pursue an export-led growth policy. Instead it is a 'policy of an autonomous rate of interest, unimpeded by international preoccupations, and a national investment programme directed to an optimum level of employment which *is twice blessed* in the sense that *it helps ourselves and our neighbors at the same time*. And *it is the simultaneous pursuit of these policies by all countries together which is capable of restoring economic health and strength* internationally, whether we measure it by the level of domestic employment or by the volume of international trade' [Keynes, 1936, p. 349, italics added].

From 1982 to 1986, the Reagan Administration unwittingly pursued this Keynesian truth by increasing military (public domestic investment) spending and cutting taxes to stimulate consumption without worrying about the economic repercussions on the United States' trade balance or federal government deficits. By mid-1982, the Federal Reserve was helping the Reagan 'Keynesian' economic expansion by reducing interest rates in order to avoid a massive international debt default. As a result the US acted as the 'engine of growth' during the Reagan Administration and most of the OECD nations rapidly recovered from the (1978-1981) greatest global recession since The Great Depression of the 1930s. Unfortunately, as recovery occurred, most of the major trading partners of the US did not engage in a 'simultaneous pursuit of these policies' of increasing public spending and reducing interest rates. These other nations neither remembered nor understood Keynes's recommendation that only by the concurrent independent expansionary public investment policies of all nations could global economic health and strength be restored.[5] Instead, some of America's trading partners took advantage of Reagan's 'Keynesian' policy, which stimulated US demand for imports, to pursue an export-led growth policy which, though initially successful, ultimately resulted in severe economic problems for countries such as Japan, much of Western Europe and the nations of southeast Asia.

Until we understand Keynes's *General Theory* lessons in an open economy context, we are doomed to repeat the past errors encouraged by 'the inadequacy of the *theoretical* foundations of the *laissez-faire* doctrine' [Keynes, 1936, p. 339] and by 'orthodox economists whose common sense has been insufficient to check their faulty logic' [Keynes, 1936, p. 349]. In a laissez-faire environment, orthodox economists assume global full employment automatically follows, so that free trade must increase the global wealth of nations by reducing each nation's aggregate supply constraints through the law of comparative advantage.[6]

In a passage that is amazingly prescient for the economic environment since Breton Woods, Keynes warns that the law of comparative advantage is only

applicable *after* all nations have implemented domestic demand management policies assuring full employment. Whenever nations operate under a laissez-faire mentality that produces significant global unemployment, then

> if a rich, old country were to neglect the struggle for markets its prosperity would droop and fail. But if [all] nations can learn to provide themselves with full employment by their domestic policy … there need be no important economic forces calculated to set the interest of one country against that of its neighbours. There would still be room for the international division of labour and for international lending in appropriate conditions. But there would no longer be a pressing motive why one country need force its wares on another or repulse the offerings of its neighbour, not because this was necessary to enable it to pay for what it wished to purchase, but with the express object of upsetting equilibrium in the balance of payments so as to develop a balance of trade in its own favour [i.e., export-led growth policy]. *International trade would cease to be what it is, namely, a desperate expedient to maintain employment at home by forcing sales on foreign markets and restricting purchases, which, if successful, will merely shift the problem of unemployment to the neighbour which is worsted in the struggle*, but a willing and unimpeded exchange of goods and services in conditions of mutual advantage [Keynes, pp. 382–3 italics added].

Unfortunately, most governments have been misled by mainstream economists to believe that free trade *per se* is job-creating globally. Keynes's *General Theory* suggests otherwise. The post-Bretton Woods international payments system with flexible exchange rates has created perverse incentives that set trading partner against trading partner to perpetuate a world of slow growth (if not stagnation). Generalizing Keynes's *General Theory* to an open economy provides a rationale for designing an international payment system that creates incentives for each nation to pursue domestic demand policies that ensure full employment without the fear of any balance of payments constraint. Only then will the gains from the law of comparative advantage become relevant.

A consistent theme throughout Keynes's *General Theory* is that classical logic has assumed away questions that are fundamental to a market-oriented, money-using economy. These problems are particularly relevant for understanding the current international payments relations that involve liquidity, persistent and growing debt obligations, and the importance of stable rather than flexible exchange rates.

An example of the sanguine classical response to Post Keynesians raising these issues is Professor Milton Friedman's response to me in our 'debate' in the literature. Friedman [1974, p. 151] stated: 'A price may be flexible…yet be relatively stable, because demand and supply are relatively stable *over time*. … [Of course] violent instability of prices in terms of a specific money would greatly reduce the usefulness of that money'. It is nice to know that as long as prices or exchange rates remain relatively stable, or 'sticky' over time, there is no harm in permitting them to be flexible. The problem arises when exchange

rates display volatility. Should there be a deliberate policy that intervenes in the market to maintain relative stability or should we allow a free market to determine the exchange rate? Keynes helped design the Bretton Woods Agreement to foster action and intervention to stabilize exchange rates and control international payment flows. Friedman sold the public on the beneficence of government inaction and the free market determination of exchange rates.

Nowhere is the difference between the Keynes view and the view of those who favor laissez-faire arrangements more evident than in regard to these questions of international capital movements and payments mechanisms and the desirability of a flexible exchange rate system. Keynes's *General Theory* analysis suggests that government monitoring and, when necessary, control of capital flows, is in society's interest. Such controls are not an infringement on the freedom of economic agents any more than the control of people's right to shout 'fire' in a crowded theater is an infringement of the individual's right of free speech.

Old Keynesians (e.g., Samuelson, Tobin, Solow) as well as New Keynesians have little to say about international capital movements and their potential detrimental effects on the balance of payments and global employment.[7] Keynes, on the other hand, recognized that large, unfettered capital flows can create serious international payments problems for nations whose current accounts would otherwise be roughly in balance. Unfortunately, in a laissez-faire system of capital markets there is no way of distinguishing between the movement of floating and speculative funds that take refuge in one nation after another in the continuous search either for speculative gains, or precautionary purposes, or for hiding from the tax collector, or laundering illegal earnings vis-à-vis funds being used to promote genuine new investment for developing the world's resources.

The international movement of speculative, precautionary, terrorist funds or illegal funds (hot money), if it becomes significantly large, can be so disruptive to the global economy as to impoverish most, if not all, nations who organize production and exchange processes on an entrepreneurial basis. Keynes warned 'Loose funds may sweep round the world disorganizing all steady business. Nothing is more certain than that the movement of capital funds must be regulated'[Keynes, 1980, p. 25]

One of the more obvious dicta that follows from Keynes's revolutionary vision of the importance of liquidity in open economies is that

> There is no country which can, in future, safely allow the flight of funds for political reasons or to evade domestic taxation or in anticipation of the owner turning refugee. Equally, there is no country that can safely receive fugitive funds which cannot safely be used for fixed investments and might turn it into a deficiency country against its will and contrary to the real facts [Keynes, 1980, p. 87].

Even in these days of global electronic communication, nations can monitor and control international capital flows only if they have the will and the necessary cooperation of other nations to do so. Monitoring and control of capital fund movements is a technical matter involving the reporting of records that are kept in the accounting system of every banking community. As long as governments have the power to tax and central bankers have the power to audit and regulate their respective domestic banking systems, large scale international capital flows can be monitored and regulated, provided there is international cooperation in this matter. As long as currency is issued only in small denominations, the physical bulkiness of moving large currency sums secretly across borders ensures that it cannot be a major threat to any capital monitoring and control policy.

In recent years, however, governments' willingness to avoid capital fund monitoring has made it easy to hide not only legally earned income and wealth from tax collectors but also profits from drug and other illegal activities and terrorist funding from law enforcement agencies. This laissez-faire attitude encourages uncivilized behavior by self-interested economic agents – and thereby imposes an important, if often neglected, real cost on society. During the 1980s flight capital drained resources from the relatively poor nations towards the richer ones, resulting in a more global inequitable redistribution of income and wealth.

Cooperation between nations in monitoring, reporting and controlling disruptive capital funds movements among nations can be readily accomplished by the international payments mechanism if some form of an international clearing union payments system of a type that as I have discussed in detail elsewhere [Davidson, 2002] were to be instituted.[8] My proposed system is developed from Keynes's Bretton Woods proposal for an international clearing union, but it does not require a Supra National Central Bank as Keynes's proposal did.

The successful implementation of my proposed international payments scheme in tandem with some rules for coordinating incomes policies among nations would ensure very inelastic expectation elasticities regarding the rates of exchange among various nations' monies. These inelastic expectations would mean that individuals will no longer be impelled to engage in disruptive international speculative and precautionary financial transactions. Thus, within a very short span of calendar time after a new payments scheme similar to the one proposed here is implemented, problems of speculative and precautionary 'hot money' flows as well as funds used to organize anti-social and illegal activities such as terrorist organizations and drug cartels and movements of income and wealth to avoid the tax collector or law enforcement officers will quickly shrink to relative insignificance.

The free world embarked on its great classical experiment of floating rates in 1973. Since 1973, there have been periodic bouts of great inflation, increasing rates of global unemployment, persistent growth in international debt, and an

increasingly inequitable international distribution of global income. Some of the rich nations got richer, while many of the poor nations became poorer at least relatively and, at times, suffered huge 'capital flight' losses.

Moreover, since 1982, one nation – the United States – has been able to take advantage of the existing international payments system to obtain a 'free lunch', that is, to run the massive perpetual trade deficits that has made the richest nation in the world also the world's largest debtor. Although residents of most other nations may resent the ability of the United States to use the international payments system to obtain this 'free lunch', they are hesitant to change a system that is heralded by modern classical economists as the only mechanism that permits everyone the freedom to choose. To be against the existing system is considered to be anti-free market and for controls by government, a particularly unpopular position in these days when State planning has apparently failed so spectacularly in Eastern Europe. In the absence of a complete collapse of the international monetary payments system, however, unless an attractive feasible alternative to the current system is put on the public agenda for discussion and development, the status quo will remain. It is an old adage in political science that 'you can't beat somebody with nobody!'

Any suggestions for reforming the international payments mechanism should build on whatever advantages the current system possesses, while providing rules to prevent any one nation from enjoying a free lunch – unless a free lunch is available to all. It is possible to provide all with a free lunch (i.e., increased global employment) if a new payments system has a built-in expansionary bias that encourages all nations to operate closer to full employment than the existing system does.

THE FACTS VERSUS THE THEORY OF FLEXIBLE EXCHANGE RATES

Because of the success of the Keynesian Revolution in stimulating domestic full employment via conscious policies between the end of World War II and the mid-1960s, the problem of wage-cost inflation became endemic to most of the developed countries of the world. Without the persistent threat of large-scale unemployment, workers and labor unions in many OECD nations became more truculent in their wage demands. By the late 1960s many developed nations were forced to pursue so-called 'stop-go' policies that generated small planned recessions to reduce the market power of labor to demand inflationary wage increases. These recessions were then followed by Keynesian policies of expansionary domestic spending, moving the economy back towards full employment until the next round of inflationary wage demands were tabled by workers. Thus the nomenclature of 'stop-go' policies.

Any nation that can pursue a successful export-led growth policy during the flexible exchange rate system that followed the Bretton Woods system breakdown in 1973 is able to export to its trading partners not only its unemployment propensities but also any inflationary tendencies. In a fixed exchange rate system, on the other hand, export-led growth policy does not provide a nation any advantage in permitting more employment and growth *with lower inflation rates* compared with Keynesian policies that stimulate domestic components of aggregate demand to achieve greater employment. Moreover, under a fixed exchange rate system, it is possible for all nations to simultaneously pursue coordinated growth oriented policies by stimulating domestic components of effective demand without necessarily running into balance of payments difficulties as long as the ratio of each nation's growth rate relative to its trading partners' growth rate is equal to the ratio of the elasticity of demand for its exports compared to the its elasticity of demand for imports (according to Thirlwall's Law[9]).

In other words, a flexible exchange rate regime guarantees that for every 'successful' economy that pursues a mercantilist trade surplus policy that can promote employment growth without significant inflation, there must be one or more offsetting 'failure' nations importing inflation and unemployment while being plagued with persistent trade deficits. For every winner on the flexible rate system, there must be one or more losers. A fixed exchange rate regime in tandem with intelligent domestic demand and incomes management policies, on the other hand, can provide entrepreneurs with profitable expansionary market opportunities in a global environment where all nations are winners. A fixed exchange rate system combined with intelligent international cooperative Keynesian policies, therefore, holds out the promise that all nations can avail themselves of a free lunch.

Since the breakdown of Bretton Woods, it has been popular to assume that freely fluctuating exchange rates in a laissez-faire market system are efficient. Every well-trained mainstream economist, whose work is logically consistent with a Walrasian, Arrow-Debreu microfoundation, 'knows' that the beneficial effects of a freely flexible exchange rate include (1) the impossibility of any one country running a persistent balance of payments deficit and (2) each nation is able to pursue monetary and fiscal policies for full employment without inflation independent of what is occurring in its trading partners[10].

The facts since the breakup of Bretton Woods, however, do not appear to be consistent with these Panglossian promises. Between 1980 and 2005, when the United States (perhaps unwittingly) took over as the engine of growth role for the global community, it has run persistent annual trade deficits. The persistent balance of payments deficits that the United States has experienced in the past quarter of a century, has permitted the economic growth miracles of (1) Japan and Germany in the 1980s, (2) the newly industrialized Pacific Rim countries

in most of the 1990s and (3) China, India and Ireland in the early years of the 21st century. The US trade deficit has permitted these nations to run high export-led growth rates during the years suggested in the last century. In that sense the United States deficits have been the engine of growth for the successful nations of the rest of the world.

Meanwhile, in the last quarter century, the floating exchange rate system, rather than encouraging the flow of capital from capital-rich developed nations to capital-poor less developed economies, has often encouraged capital flight in the opposite direction, thereby draining resources from the relatively poor nations of Africa and Latin America towards the richer ones. The result has been a growing global inequitable redistribution of income and wealth, thereby increasing the immiseration of the majority of the people on this planet.

In sum, then, during this period of floating rates since the 1970s, the world has not achieved the sate of economic bliss promised by classical theory. If anything, the economic situation for a majority of the world's population has deteriorated.

In a world operated according to classical axioms, export-led growth should be no more desirable in terms of generating employment without inflation than internally generated demand-led growth. Classical economics assumes that the economy will track the long-run full employment growth trend no matter what the primary source of demand growth. Yet the facts of the 1980s demonstrate that all 'successful' economies tend to pursue export-led growth rather than domestic demand induced expansion. Nations such as Germany, Japan, Taiwan, Singapore, Hong Kong, and South Korea were not only applauded for their economic miracles by leading Monetarist and Old and New Keynesian scholars, but they were held up as shining examples of the proper functioning of a classical economy operating free from oppressive government intervention. Yet there is nothing in classical theory that justifies relying primarily on export-led growth!

THIRLWALL'S LAW

Professor A.P. Thirlwall (1979) has developed Keynes's multiplier mechanism into a demand-driven model of economic growth that does not make the classical presumption of continuous global full employment. In his formulation, Thirlwall posits traditional export and import demand functions:

$$X_a = (P_d/P_f)^z Y^{erw} \tag{1}$$

$$M_a = (P_d/P_f)^u Y^{ea} \tag{2}$$

where X_a and M_a are exports from nation A and imports into A during a period, (P_d/P_f) is the ratio of domestic prices to foreign prices expressed in terms of the domestic currency of A, z is the price elasticity of demand for imports, Y is income for each region, e_a is A's income elasticity of demand for imports, and e_{rw} is the rest of the world's income elasticity of demand for A's exports. If either z and u are small and/or relative prices do not change significantly, then, as a first approximation, one can ignore substitution effects and concentrate on income effects. Taking the natural logs of equations (1) and (2) and ignoring substitution effects, one obtains Thirlwall's Law of the growth of income, which is consistent with an unchanged trade balance, as

$$y_a = x/e_a \tag{3}$$

where y_a is the rate of growth of Nation A's GNP, x is the rate of growth of A's exports, and e_a is A's income elasticity of demand for imports. Since the growth of exports for A depends primarily on the rest of the world's growth in income (y_{rw}) and the world's income elasticity of demand for A's exports (e_{rw}), i.e.,

$$x = (e_{rw})(y_{rw}) \tag{4}$$

equation (3) can be written as

$$y_a = [e_{rw}\, y_{rw}]/e_a \tag{5}$$

The rate of growth that a nation can maintain without running into a balance of payments problem depends on the rest of the world's growth and the relevant income elasticities for imports and exports. If the growth of imports is to exactly equal the growth in the value of exports,

$$e_{rw}y_{rw} = y_a e_a \tag{6}$$

then,

$$[y_a/y_{rw}] = e_{rw}/e_a \tag{7}$$

i.e. the ratio of the growth of income in nation A compared to growth in income in the rest of the world is equal to the ratio of the income elasticity of demand for A's exports by the rest of the world to A's income elasticity of demand for imports. Thus, for example, if $e_{rw}/e_a < 1$, and if growth in A is constrained by the need to maintain balance of payments equilibrium, then nation A is condemned to grow at a slower rate than the rest of the world.

If, for example, less developed nations (LDCs) of the world have a comparative advantage in exports of raw materials and other basic commodities, for which Engel's curves suggest that the developed world will have a low income elasticity of demand, while the LDCs have a high income elasticity of demand for the manufactured products and services of the developed world, then for most LDCs

$$[e_{rw}/e_{ldc}] < 1 \tag{8}$$

Accordingly, if economic development and balance of payments equilibrium is left to the free market, the LDCs are condemned to relative poverty, and the global inequality of income will become larger over time.

Moreover, if the rate of population growth in the LDCs (p_{ldc}) is greater than the rate of population growth in the developed world (p_{dw}), that is, if $p_{ldc} > p_{dw}$, then the rate of growth of GNP per capita of the LDCs will experience a greater relative decline to the standard of living of the developed world, i.e.,

$$[y_{ldc}/p_{ldc}] < < [y_{dw}/p_{dw}] \tag{9}$$

In the absence of Keynesian policies to stimulate growth, the long-term growth rate of the developed world taken as a whole tends to be in the 1–2.5 percent range. As long as the developed world's population growth is less than its long-term growth rate, however, these nations can still enjoy a rising living standard.

If, however, we accept the reasonable values for the parameters implied in inequality (9), then since $y_{ldc} < y_{dw}$, we get $1 < y_{dw} < 2.5$; a dreary prognostication for the global economy emerges. As long as the world permits the free market to determine the balance of payments constraint on each nation, then a shrinking proportion of the world's population may continue to get richer (or at least hold their own), while a growing proportion of the earth's population is likely to become poorer, if not absolutely then at least relatively. Moreover, the slower the rate of growth in income of the rich, the more rapidly the poor are likely to sink into poverty.

Thus, there is an obvious case to explore if there are some policy interventions that can be developed to prevent market determined balance of payments constraints from condemning the majority of the world's population to increasing poverty. Only if the rich can achieve the historically high real rates of growth experienced in the first 25 years since World War II (where Keynesian rather than free market policies were actively pursued domestically and internationally by the developed world) can we hope to significantly improve the economic lot of the poorer nations of the world.

Since the US has apparently not been significantly constrained by payments deficits in the last quarter century, equation (8) can be interpreted in a different

light for the US Given the US rate of growth under Reagan and since, then, if one assumes the import and export income elasticities of demand (e_{rw} and e_a) are fixed, then solving equation (8) for y_{rw} yields the income growth that would have been required of the US's trading partners in order to eliminate the US trade deficit. Alternatively, if y_{rw} is presumed unchanged, then solving for e_{rw} would give the required income elasticity necessary to avoid a US trade deficit.

Consequently, Thirlwall's Law analysis demonstrates that international financial payment imbalances can have severe real consequences, i.e., money is never neutral in an open economy. This suggests that the nations of the world should cooperate in developing an international monetary and payments system similar to Keynes's clearing union and/or my international clearing union mechanism mentioned above that permits and encourages the globalized economy of the 21st century to grow at a close to full employment rate.

NOTES

1. John Maynard Keynes, *The General Theory of Employment Interest and Money* (New York: Harcourt, 1936). All page references to passages from this book will be cited in the text of this chapter accompanying the relevant quote or discussion. References to any other writings of Keynes or any other will appear as endnotes.
2. Nations with banking institutions which make it difficult for foreign authorities to obtain information regarding bank accounts held by their residents are likely to encourage the influx of funds trying to escape national tax collectors, criminal investigators, and the central banks of nations that try to limit capital outflows. Thus, it is not surprising, that often exchange rates reflect speculative, and flight capital flows rather than purchasing power parities.
3. For example, in 1977 the Carter Administration's attempted to 'talk down the dollar'. In the Spring of 1993, Secretary of Treasury Bentsen tried to talk up the yen. In January 1994, the New York Times quoted Secretary Bentsen as saying that 'allowing the yen to decline would not be an acceptable way for Japan to try to escape from its recession'.
4. Most mainstream economists were appalled by President Reagan's boasts regarding the higher dollar that was achieved in the early years of his Administration.
5. Even as this is being written nations are still ignoring this Keynesian 'truth' to the detriment of over 38 million unemployed people in the OECD nations and many more in Eastern Europe and the former Soviet Union.
6. In this matter, Keynes pointed out, 'the [orthodox] faculty of economists prove to have been guilty of presumptuous error' [Keynes, 1936, p. 339].
7. As I point out in my book, *Post Keynesian Macroeconomic Theory* [Davidson, 1994], both Old and New Keynesian analysis is based on the restrictive classical axioms that Keynes threw out in developing his *General Theory*. It is no wonder therefore that these 'Keynesians' subscribe to the classical view of international trade – if they think about it at all.
8. To argue, from the outset, that international cooperation in sharing records and helping enforce capital flows cannot be achieved, is unduly pessimistic. It paints a picture of the human condition where nations were willing to cooperate in world wars at a cost of the lives of a large portion of their youth in the 20th century, but unwilling to cooperate even if it costs the recipient nations a 'fast buck' in the 21st century.
9. For a discussion of Thirlwall's Law see Section III *infra*.
10. In 1968, Professor Harry Johnson wrote [in *The Times* of London, 12/9] 'the basic argument for floating exchange rates is so simple that most people have considerable difficulty in un-

derstanding it. … a floating exchange rate would save a country from having to reverse its full employment policies because they lead to inflation and deficit'.

REFERENCES

Davidson, P. (1994), *Post Keynesian Macroeconomic Theory*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Davidson, P. (2002), *Financial Markets, Money and the Real World*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Friedman, M. (1974), 'A Response To His Critics', in R.G. Gordon (ed.), *Milton Friedman's Monetary Framework: A Debate With His Critics*, Chicago: Chicago University Press.

Keynes, J.M. (1936), *The General Theory of Employment Interest and Money*, New York: Harcourt.

Keynes, J.M. (1980), *The Collected Writings of John Maynard Keynes*, 25 volumes, ed. D. Moggridge, London: Macmillan.

Thirlwall, A.P. (1979), 'The Balance of Payments Constraint As An Explanation of International Growth Rate Differences', *Banca Nazionale del Lavoro Quarterly Review*, **128**.

5. Cycles, aggregate demand, and growth

Miguel A. León-Ledesma

INTRODUCTION

Growth theory has focused on the causes of increases in the levels of per capita income from a perspective that assumes that, generally, the business cycle has no role to play. This assumption eliminates any influence of cyclical behaviour stemming from aggregate demand and nominal shocks on the long-run performance of nations. That is, growth theory is a theory of *potential output growth*. This contrasts with the more policy-oriented and popular view that good macroeconomic management is a pre-condition for healthy and sustainable growth in the long run. The question then arises as to whether demand shocks and other determinants of the business cycle do have a role to play in determining the level of potential output.

The relation between cycles and growth is not a new issue in macroeconomics, but revived interest on it arose as a consequence of the development of the endogenous growth models of Romer (1986 and 1990), Lucas (1988) and Aghion and Howitt (1992). Back in the 1960s and 1970s, this relation was also tackled within Keynesian frameworks by authors such as Kaldor (1966 and 1970) and Thirlwall (1979). The Real Business Cycle (RBC) literature, on the other hand, has continued to assume that business cycles do not affect potential output, hence eliminating non-linearities arising in the decomposition of shocks.[1] Recently, with the development of new datasets and econometric techniques, authors such as Malley and Muscatelli (1999) and Pedersen (2003) have attempted to unveil statistical relations between cycles and productivity.

In a recent series of papers by Tony Thirlwall and myself, see León-Ledesma and Thirlwall (2000 and 2002), we claim that the natural rate of growth is endogenous to changes in the actual rate of growth of output.[2] That is, continued expansions or contractions can lead to changes in the potential growth rate of output through their impacts on productivity growth and the supply of labour. This idea links two strands of the literature back together. On the one hand the Keynesian/Kaldorian tradition where growth is demand-led (see McCombie and Thirlwall, 1994) and, on the other hand, the cycles and growth literature derived from endogenous growth models. Given the focus of León-Ledesma and Thirl-

wall (2002) on the first strand of the literature, here I will focus on the second, and contrast them where the two overlap.

The issue has obvious implications for macroeconomic policy because, if potential output depends on the state of the cycle, so will *output gaps*. Anti-inflationary policies based on output gap indicators within Phillips Curve with Taylor Rule frameworks would need to look into these issues. This is of special relevance within the Euro area given the way macroeconomic policies are designed and constrained with a strong monetary policy-dominance. A close look into these issues would require an analysis of the role that labour market and goods market rigidities play in the direction and strength of the relation between cycles and growth.

Hence, in this chapter I will discuss the possible relations between cycles, aggregate demand shocks and growth, focusing, as mentioned earlier, on the endogenous growth literature and its links with the Kaldorian tradition represented by Tony Thirlwall. I will also present a new and simple method for analyzing the relationship between cycles and growth that endogenizes the impact of cycles on the definition of trend output. I apply this method to analyze the cases of the US, Germany and the UK. My findings show that cycles do seem to have a strong impact on trend output but this impact is different for the three economies in question. As such, this is just a first step to integrate these issues and should be seen more as prospective analysis than any final answer, which should come from further and more detailed research.

The chapter is organized as follows. Next section presents the different mechanisms put forward in the literature linking cycles and growth. Section 3 describes the econometric methodology used in this study. Section 4 comments on the results, and the final section draws some conclusions.

GROWTH AND CYCLES: THE MECHANISMS

There are several mechanisms that can link business cycles and growth (see Saint-Paul, 1997 and Aghion and Howitt, 1998 for surveys of this topic). Broadly, we can classify them into two groups. The first emphasizes the positive impact of cycle upturns on productivity. The second emphasizes the positive impact of recessions on productivity.

The idea that cyclical booms can affect long-run growth has roots on the *learning-by-doing* (LBD) idea of Arrow (1962) that has been partially taken up by endogenous growth models. Expansions due to demand or supply shocks would increase the size of the market, inducing division of labour and investment in capital, which will in turn generate a learning curve enhancing productivity and long-run growth. Modern models such as Stadler (1990) have focused on the impact of expansions on R&D activity. If firms face financial

constraints, boom periods will allow firms to finance R&D through retained profits. This pro-cyclicality of R&D, emphasized also by Stiglitz (1993), would induce an impact of demand shocks on long-run productivity. There is no need, however, to resort to explicit R&D for generating this mechanism as we know that R&D is usually carried out by large firms, or small firms that are highly dependent on large ones. If financing constraints à la Fazzari et al (1988) are predominant, expansionary periods would induce higher investment in capital. As capital embodies technical progress or is complementary with human capital, firms' productivities would increase without explicit R&D. There is, however, another related link. As booms expand the size of the market, the scope for division of labour and *roundabaoutness* increases productivity in the way that was already pointed out by Young (1928).

These mechanisms are capable of opening up avenues through which aggregate demand shocks, either real or nominal, can influence long-run growth. Empirical studies of RBCs have only been able to attribute to technology shocks less than 30% of output variations at medium-run horizons (see Christiano et al, 2003 and Galí, 1999). One is tempted to infer that a larger role in fluctuations is played by aggregate demand and hence, demand factors may have a role to play in determining potential output. This counters the modelling approach of the vast majority of growth theory. The influence of demand on growth through its impact on productivity is a theme that has been largely the focus of cumulative causation models of growth. These models, based on Kaldor (1970) and Dixon and Thirlwall (1975), emphasize the role that external demand has on expanding output and initiating cumulative processes of productivity expansion. The idea is that demand expansion leads to increased productivity through the Verdoorn effect, which increases price competitiveness, and in turn, export expansion. Exports are the only component of aggregate demand whose expansion would not cause balance of payments constraints.

Recently León-Ledesma (2002) has presented a model along these lines that incorporates the role of R&D, LBD and technology diffusion. The model is able to generate a rich set of dynamic paths for relative productivities, and is compatible with both diverging and converging productivity levels across countries. Models like this provide a link between demand and growth through induced productivity that resembles those mentioned earlier. In fact, the central mechanism of cumulative causation is the Verdoorn effect, which relates output and productivity. Given the black-box nature of this effect, cumulative causation models could well embrace micro-founded explanations of productivity growth with credit constraints such as those presented in Stadler (1990). The important message in both cases is that sustained periods of expansion can lead to productivity gains that increase the natural rate of growth. If cyclical fluctuations are mostly demand-driven then the link between demand, cycles and growth

that arises both in endogenous growth and cumulative growth models would be of the same nature.

On the other hand, recent models of endogenous growth through 'creative destruction' have pointed to the possible positive impact of recessions on long-run growth. In essence, the approach takes an evolutionary selection view in which recessions clean industries from their inefficient units, increasing average productivity (Caballero and Hammour, 1994). Also, the reorganization of activities within firms usually takes place during recessions, as the opportunity cost of restructuring is lower, given that the cost of lost production and asset values is lower (Hall, 1991). These mechanisms would then imply that long-run trend growth would be positively affected by recessions.

Empirical evidence is still scarce[3] and mainly inconclusive. In addition, most of the evidence focuses on the US experience, with EU countries left aside. The evidence using disaggregated data points to the important role of the opportunity cost mechanism. However, disaggregated data may have the disadvantage that they do not account for aggregate effects on productivity stemming from externalities and aggregate division of labour. Also, time series dimensions for this kind of data are usually small and hence they tend to work with particular periods of expansion or recession rather than providing a more general view of the impact of cycles. Finally, there is little evidence that tries to identify the importance of different possible underlying links between cycles and growth.

EMPIRICAL METHODOLOGY

As an initial attempt to analyze these relations, we propose the use of non-linear univariate time series models that allow the equilibrium levels of the variables to vary depending on a threshold variable. The idea is the following. We can estimate an autoregressive model for output in which the equilibrium or trend level of output depends on the previous state of the cycle. The obvious candidate is the family of so-called Threshold Autoregressive models (TAR) due to Tong (1983). We allow output to have a different trend component and persistence pattern depending on whether the cycle is above or below a certain threshold and, hence, see how that equilibrium is affected by the cycle. This is related to the existing evidence on the asymmetric behaviour of output as first pointed out by Nefcti (1984).

Specifically, if output behaves asymmetrically, we can use the following TAR (Threshold Autoregression) representation (Caner and Hansen, 2001):

$$\Delta y_t = \theta_1' y_{t-1} 1_{\{z_{t-1}<\lambda\}} + \theta_2' y_{t-1} 1_{\{z_{t-1}\geq\lambda\}} + \sum_{j=1}^{p} \gamma_j \Delta_{t-j} + \zeta_t, \qquad (1)$$

where $y_{t-1} = (1\ t\ x_{t-1})$, x_t is output and $1_{\{\cdot\}}$ is the indicator function that takes the value of 1 if z_{t-1} is higher or lower than a threshold λ, and 0 otherwise. The variable z_t is any stationary variable that would determine the change of regime or, in our case, the cycle. For our purposes, we can set z_t in several different ways. There is no obvious candidate to measure cycles, but here we will work with two simple definitions that may be subject to further refinement:

- Set the cycle indicator as $z_t = x_t - x_{t-m}$, that is, the increase in output over a period of m quarters.
- Set the cycle indicator as the difference between actual output and a smoothed estimate of output such as the Hodrick–Prescott (HP) filter, that is, the output gap, $z_t = (x - hpx)_{t-m}$.[4]

We assume that output may have a different behaviour depending on whether (i) past *changes* in output have been higher or lower than a certain threshold λ; (ii) the output gap is higher or lower than a certain threshold λ; (iii) the output gap is higher or lower than zero. We call these models Δy, $HP_{gap}1$ and $HP_{gap}2$ respectively. The first model is a momentum-TAR model or M-TAR as in Enders and Granger (1998). The lag length m for the changes in output and the output gap will be data determined as will be the search for the optimal threshold λ. Finally, the parameter vectors θ_1 and θ_2 can be partitioned as

$$\theta_1 = \begin{pmatrix} \mu_1 \\ \tau_1 \\ \rho_1 \end{pmatrix}, \quad \theta_2 = \begin{pmatrix} \mu_2 \\ \tau_2 \\ \rho_2 \end{pmatrix}$$

The choice of the threshold λ could be simply made on an a priori basis, such as setting $\lambda = 0$ as in the third model. However, this would be a biased estimate of the threshold, if asymmetric adjustment exists, and a subjective measure. In order to search for the optimal threshold in models 1 and 2, we follow Chan (1993) and find λ as the value of z_t that minimizes the residual sum of squares of the OLS estimation of (1).[5]

Our aim is to test for different behaviour of output depending on the state of the cycle, i.e. asymmetry. In order to test for the existence of asymmetry in the adjustment under both regimes we test the null hypothesis $H_0 : \theta_1 = \theta_2$ on the OLS estimation of (1), making use of the Wald statistic (W) proposed in Caner and Hansen (2001). Finally, we also choose m to minimize the residual sum of squares. Given that the Wald test of asymmetry is a monotonic function of the residual variance, we choose m as the value which maximizes the Wald test of asymmetry. Output, however, may have a unit root or near unit root. This is confirmed when we apply three different unit root tests to our data for the US, UK and Germany for the period ranging from 1960 Q1 to 2001 Q2. Table 5.1

Table 5.1 Unit root tests on output

		ADF		*KPSS*		*Elliott DFGLS*$_u$	
	Lag	Constant	Trend	Constant	Trend	Constant	Trend
Germany	4	–1.262	–2.536	3.325	0.342	–0.603	**–3.431**
UK	3	–0.390	–2.201	3.349	0.250	–0.255	–2.732
US	2	–1.235	–3.333	3.365	0.329	–0.953	**–3.323**

Note: Bold indicates the rejection of the null of a unit root for the ADF and DFGLS tests and the acceptance of the null of stationarity for the KPSS test at the 5% level.

reports the results of applying the ADF, KPSS and Elliott (1999)-DFGLS tests. Just two out of the eighteen tests reject the null of a unit root in the ADF and DFGLS tests or accept the null of stationarity in the KPSS test. This poses non-trivial problems when testing for asymmetry. We hence follow Caner and Hansen (2001) and test simultaneously for asymmetry and unit roots. Stationarity would imply rejecting H_0: $\rho_1 = \rho_2 = 0$, and we also make use of two Wald statistics (R1 and R2).

The procedure we follow to test simultaneously for asymmetry and unit roots implies first estimating a baseline model for the linear ADF regression to determine the lag augmentation of the DF regression using general-to-specific techniques as used for finding the augmentation lag in the tests reported in Table 5.1. We then select the threshold by minimising the residual sum of squares of (1) as mentioned earlier and fit the TAR model by OLS for every value of m. We choose the m that minimizes the residual sum of squares for all values of m.

Given that the asymptotic null distribution of the asymmetry test (W) is non-standard, Caner and Hansen (2001) recommend the use of bootstrap methods to obtain p-values. In a Monte Carlo experiment they show that the power and size of the test does not crucially depend on whether we impose a unit root. Hence, we obtained p-values by carrying out 1,000 iterations of the unconstrained asymmetry test, i.e. not imposing the existence of a unit root. Finally, the unit root hypothesis involves testing for H_0: $\rho_1 = \rho_2 = 0$. There are two possible alternatives: H_1: $\rho_1 < 0$ *and* $\rho_2 < 0$ and

$$H_2 : \begin{cases} \rho_1 < 0 \text{ and } \rho_2 = 0 \\ \text{or} \\ \rho_1 < 0 \text{ and } \rho_2 = 0 \end{cases}$$

The first alternative corresponds to the stationary case, whilst the second implies stationarity in only one of the regimes, which implies overall non-stationarity

but a different behaviour from the classic unit root. Caner and Hansen (2001) develop asymptotic theory for the distribution of this unit root test. However, for finite samples they recommend the use of bootstrapping. As the distribution of the test statistic will depend on whether or not a threshold effect exists, p-values obtained through the bootstrap are not unique. We hence obtained the bootstrapped p-values from 1,000 iterations under the hypothesis that the threshold is not identified (R1) and under the hypothesis that it is identified (R2). These two tests have substantially more power than the ADF test as threshold effects become more important. In order to discriminate between the two alternatives in H_2, Caner and Hansen (2001) recommend looking at the t-ratios of ρ_1 and ρ_2.

Finally, given our interest in the impact of cycles on trend growth, we also carried out an F-test for equality of the trend parameters H_0: $\tau_1 = \tau_2$. This test, however, should be taken with more caution as we relied on standard F-distributions to test for significance.

RESULTS

Table 5.2 reports the results of the asymmetry and unit root tests together with their bootstrapped p-values. The table also reports the lag of the threshold variable *m* and the optimal threshold for models Δy and $HP_{gap}1$. Notice that, as R1 assumes no identified threshold under the null, its value is the same for the three models. The model was estimated for three major economies, the US, the UK and Germany using quarterly data for the 41 years ranging from 1960:1 to 2001:2. The period is long enough so as to capture several episodes of booms and recessions and hence increasing the precision of our estimates of the asymmetry effects.

The first thing to notice in the results is the strong asymmetry present in the data judging by the values of the W-test. Only in two cases we can accept the null of linearity, and only at the 10% level, not at the 5%. This asymmetry seems to be stronger for Germany, followed by the US and the UK in that order. The F-test on the trend parameter produces mixed results. Only in the cases of the US and Germany does trend asymmetry seem to be contributing to the overall asymmetry of the model. In the case of the UK, the trend component of output does not seem to change with changes in the business cycle.

The unit root tests are somewhat surprising. Once we account for the possible presence of asymmetries, it appears that output is stationary around a trend. This contradicts much of the evidence on real output non-stationarity that considers a linear representation under the alternative.[6] For the case of the US and Germany, where asymmetry is stronger, the rejection of the null under the identified threshold test R2 is stronger, as expected.

Table 5.2 Tests of threshold effects tests for output

Country	US			GER			UK		
	Δy	$HP_{gap}1$	$HP_{gap}2$	Δy	$HP_{gap}1$	$HP_{gap}2$	Δy	$HP_{gap}1$	$HP_{gap}2$
λ	0.019	0.041	0	0.075	0.031	0	0.012	0.033	0
m	4	3	4	5	2	3	3	2	1
	19.656[a]	**11.069**	7.920	**32.307**	**13.874**	**12.247**	**9.278**	**11.180**	9.430
W *(p-value)*	**(0.004)**	**(0.030)**	(0.070)	**(0.000)**	**(0.028)**	**(0.016)**	**(0.038)**	**(0.042)**	(0.076)
F-test trend	**3.991**	0.026	**5.036**	**21.994**	0.773	**4.021**	0.035	1.668	0.644
(p-value)	**(0.047)**	(0.871)	**(0.026)**	**(0.000)**	(0.381)	**(0.041)**	(0.852)	(0.198)	(0.424)
R1	**11.672**	**11.672**	**11.672**	**6.934**	**6.934**	**6.934**	**9.301**	**9.301**	**9.301**
(p-value)	**(0.001)**	**(0.001)**	**(0.001)**	**(0.016)**	**(0.016)**	**(0.016)**	**(0.004)**	**(0.004)**	**(0.004)**
R2	**23.357**	**16.151**	**15.221**	**32.477**	**11.451**	**11.126**	**9.311**	**8.741**	5.299
(p-value)	**(0.002)**	**(0.007)**	**(0.018)**	**(0.001)**	**(0.008)**	**(0.022)**	**(0.016)**	**(0.005)**	(0.094)
Lag	2	2	2	4	4	4	3	3	3

Note: [a] Bold indicates rejection of the null of no asymmetry or unit root at the 5% level. P-values are obtained through the bootstrapping technique of Caner and Hansen (2001), except for the F-test on the trend.

Tables 5.3–5 present the results of the estimated parameters when we assume asymmetry for each country and model considered. There are general patterns in the three models and, although models $HP_{gap}1$ and 2 are the closest two, the choice of the cycle indicator does not affect strongly the pattern of cyclical behaviour of real output. To ease interpretation of the results, it is worth reminding that parameters with the subscript 1 correspond to expansionary phases of the cycle. In all three countries the velocity of adjustment to the trend component (ρ) is higher in recessions than in booms. This is a consequence of the fact that usually expansions last longer than contractions. In many cases the series appear to be non-stationary during expansions and stationary in recessions. The exception is Germany in the Δy model. However, the large intercept found for that case casts doubts about the results for that particular model.[7]

When analyzing the parameters of the trend variable we find interesting results. First, for the UK there is practically no difference between expansions and recessions as we would expect from the previous tests. This indicates that the impact of cycles on growth in the UK is either very limited or not possible to capture using this technique. For the US there are substantial differences, especially in models Δy and $HP_{gap}2$. In this case, the trend output after a recession seems to have a higher slope. This supports a stronger impact of the opportunity cost kind of effect as opposed to LBD effects. Finally, for Germany,

Table 5.3 Estimates of the threshold model: US

	Δy	$HP_{gap}1$	$HP_{gap}2$
μ_1	**0.221**[a]	**0.232**	0.177
	(2.923)	**(1.817)**	(1.572)
μ_2	**0.557**	**0.288**	**0.320**
	(3.995)	**(3.691)**	**(3.673)**
ρ_1	**–0.061**	**–0.066**	–0.049
	(–2.812)	**–1.796)**	(–1.520)
ρ_2	**–0.157**	**–0.080**	**–0.089**
	(–3.931)	**(–3.595)**	**(–3.593)**
τ_1	**4.555e–04**	**5.403e–04**	3.608e–04
	(2.700)	**(1.721)**	(1.378)
τ_2	**0.001**	**5.987e–04**	**6.771e–04**
	(3.671)	**(3.419)**	**(3.480)**

Note: [a] Bold indicates rejection of the null of no significance at the 5% level.

Table 5.4 Estimates of the threshold model: UK

	Δy	$HP_{gap}1$	$HP_{gap}2$
μ_1	**0.291**	–0.364	–0.154
	(2.399)	(–0.612)	(–0.224)
μ_2	**0.328**	**0.236**	**0.239**
	(1.964)	**(2.303)**	**(2.343)**
ρ_1	**–0.077**	0.099	0.043
	(–2.397)	(0.628)	(0.234)
ρ_2	**–0.083**	**–0.061**	**–0.062**
	(–1.888)	**(–2.253)**	**(–2.290)**
τ_1	**4.562e–04**	–6.761e–04	–6.761e–04
	(2.503)	(–0.536)	(–0.536)
τ_2	4.979e–04	**3.397e–04**	3.397e–04
	(1.562)	**(2.210)**	(1.210)

Note: [a] Bold indicates rejection of the null of no significance at the 5% level.

Table 5.5 Estimates of the threshold model: Germany

	Δy	$HP_{gap}1$	$HP_{gap}2$
μ_1	**2.040**	0.304	0.032
	(5.135)	(1.373)	(0.214)
μ_2	**0.209**	**0.214**	**0.327**
	(2.613)	**(2.544)**	**(3.440)**
ρ_1	**–0.569**	–0.084	–0.009
	(–5.119)	(–1.360)	(–0.209)
ρ_2	**–0.056**	**–0.057**	**–0.088**
	(–2.503)	**(–2.437)**	**(–3.329)**
τ_1	**0.004**	**7.838e–04**	**0.002**
	(5.185)	**(1.740)**	**(3.221)**
τ_2	**3.533e–04**	**3.649e–04**	**5.609e–04**
	(2.244)	**(2.176)**	**(2.995)**

Note: [a] Bold indicates rejection of the null of no significance at the 5% level.

the opposite results appear. In this case the trend output growth is higher after an expansion than after a recession.

Our results, hence, are not able to determine uniquely the direction of the link between cycles and growth. Nevertheless, an important fact is that LBD effects appear to outweigh cleansing or opportunity cost effects in the economy where labour markets are acknowledged to be more rigid. The opposite happens with the US, the economy with the most flexible labour markets. Our simple approach is not able to test directly this proposition, but points to a strong role of labour market institutions on the relation between cycles and growth. A lower degree of wage adjustment to business cycles increases the adjustment in quantities and/or productivity, which explains the strong link between cycles and trend growth in Germany, as opposed to the US and UK. Another important aspect that could be influencing this relationship is financial markets and banking systems that may influence the degree of credit constraints faced by firms. A bank-based system such as the German one may lead to information asymmetries being reflected in credit rationing rather than stock prices or interest rates, which could also contribute to the pro-cyclical behaviour of potential output.

CONCLUSIONS

Textbook growth theory models *potential output growth*, leaving no role for aggregate demand and other cyclical shocks on long-run growth. This contrasts with the more policy-oriented and popular view that good macroeconomic management is a pre-condition for healthy and sustainable growth in the long run. We have attempted to contribute to the question of whether demand shocks and other determinants of the business cycle do have a role to play in determining the level of potential output.

We have done so by fitting a Threshold Autoregressive Model to output for the US, Germany and the UK. In this model, the trend or equilibrium level of real output is allowed to change, depending on the state of the business cycle. Using different definitions of cycle, we arrived at similar conclusions. Our findings show a strong asymmetry of real output. Real output also seems to be stationary during recessions. The impact of the business cycle on the trend component of output differs for each economy. For the UK, we did not find a different behaviour depending on the cycle. For the US, we found that trend output increases after recessions. The opposite happens to Germany, where trend output growth appears to be higher after an expansion. These results point to labour markets and financial systems as important determinants of the relation between cycles and growth, which is a topic open for further research. Research in this direction can have important implications for the design of stabilization policies to promote growth.

An important conclusion, as pointed out by Tony Thirlwall throughout several of his contributions to growth and development theory, is that the role of demand on long-run growth cannot be ignored. This is not to say that growth is purely demand-determined, but that the interaction between the demand and supply structures and characteristics of an economy are key factors for understanding why growth rates differ in the long run.

NOTES

1. See Christiano et al (2003) and Galí (1999).
2. See the criticism by Boggio and Seravalli (2002) and our reply in León-Ledesma and Thirlwall (2002) where we close the issue.
3. See, for instance, Baily et al (2001), Malley et al (2005), Malley and Muscatelli (1999) and Pedersen (2003).
4. We also carried out the rest of estimates using a Band-pass Filter (BP), but the results were very similar to those using the HP filter and are not reported here.
5. In practice, we eliminated the highest and lowest 10% values of z_t.
6. In his seminal paper on structural breaks and unit roots, though, Perron (1989) finds that output is stationary around a breaking trend.
7. In fact, in this model only 11% of the observations belonged to the expansionary periods. This is a consequence of the high threshold value (λ). We could have trimmed the search for λ further, but preferred to keep this result for comparison purposes.

REFERENCES

Aghion, P. and Howitt, P. (1992), 'A Model of Growth through Creative Destruction', *Econometrica*, **60**, 323–51.

Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, Cambridge, MA: MIT Press.

Arrow, J.K. (1962), 'The Economic Implications of Learning-by-doing', *Review of Economic Studies*, **29**, 155–73.

Baily, M.N., Bartelsman, E.J. and Haltiwanger, J. (2001), 'Labour Productivity: Structural Change and Cyclical Dynamics', *Review of Economics and Statistics*, **83**(3), 420–34.

Boggio, L. and Seravalli, G. (2002), 'Is the Natural Rate of Growth Endogenous? A Comment', *BNL Quarterly Review*, **55**, 219–27.

Caballero, R. J. and Hammour, M. (1994), 'The Cleansing Effects of Recessions', *American Economic Review*, **84**, 1350–68.

Caner, M. and Hansen, B.E. (2001), 'Threshold Autoregressions With a Unit Root', *Econometrica*, **69**, 1555–96.

Chan, K.S. (1993), 'Consistency and Limiting Distribution of the Least-Squares Estimator of a Threshold Autoregressive Model', *The Annals of Statistics*, **21**, 520–33.

Christiano, L.J., Eichembaum, M. and Vigfusson, R. (2003), 'What Happens After a Technology Shock?', *NBER Working Paper*, 9819.

Dixon, R. and Thirlwall, A.P. (1975), 'A Model of Regional Growth-rate Differences on Kaldorian Lines', *Oxford Economic Papers*, **27**, 201–13.

Elliott, G. (1999), 'Efficient Tests for a Unit Root when the Initial Observation is Drawn from its Unconditional Distribution', *International Economic Review*, **40**, 767–83.
Enders, W. and Granger, C.W.J. (1998), 'Unit Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates', *Journal of Business and Economic Statistics*, **16**, 304–11.
Fazzari, S., Hubbard, R.G. and Petersen, B. (1988), 'Finance Constraints and Corporate Investment', *Brookings Papers on Economic Activity*, **1**, 141–95.
Galí, J. (1999), 'Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?', *American Economic Review*, **89**, 249–71.
Hall, R.E (1991), 'Recessions as Reorganizations', *NBER Macroeconomics Annual.*
Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the UK: An Inaugural Lecture*, Cambridge: Cambridge University Press.
Kaldor, N. (1970), 'The Case for Regional Policies', *Scottish Journal of Political Economy*, **17**, 337–48.
León-Ledesma, M.A. (2002), 'Accumulation, Innovation and Catching-up: An Extended Cumulative Growth Model', *Cambridge Journal of Economics*, **26**, 201–16.
León-Ledesma, M.A. and Thirlwall, A.P. (2000), 'Is the Natural Rate of Growth Exogenous?', *Banca Nazionale del Lavoro, Quarterly Review*, **53**, 433–46.
León-Ledesma, M.A. and Thirlwall, A.P. (2002), 'The Endogeneity of the Natural Rate of Growth', *Cambridge Journal of Economics*, **26**, 441–60.
León-Ledesma, M.A. and Thirlwall, A.P. (2002), 'The Endogeneity of the Natural Rate of Growth: a Reply to Boggio and Seravalli', *Banca Nazionale del Lavoro*, *Quarterly Review*, **55**, 228–30.
Lucas, R. (1988), 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, **22**, 3–42.
Malley, J. and Muscatelli, A. (1999), 'Business Cycles and Productivity Growth: Are Temporary Downturns Productive or Wasteful?', *Research in Economics*, **53**, 337–64.
Malley, J., Muscatelli, A. and Woitek, U. (2005), 'Real Business Cycles or Sticky Prices? The Impact of Technology Shocks on US Manufacturing', *European Economic Review*, 754–60.
McCombie, J.S.L. and Thirlwall, A.P. (1994), *Economic Growth and the Balance of Payments Constraint*, London: Macmillan.
Neftci, S.N. (1984), 'Are Economic Time Series Asymmetric Over the Business Cycles?', *Journal of Political Economy*, **85**, 281–91.
Pedersen, T.M. (2003), 'International Evidence on the Connection Between Business Cycles and Economic Growth', *Journal of Macroeconomics*, **25**, 255–75.
Perron, P. (1989), 'The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis', *Econometrica*, **57**(6), 1361–401.
Romer, P.M. (1986), 'Increasing Returns and Long Run Growth', *Journal of Political Economy*, **94**, 1002–1037.
Romer, P.M. (1990), 'Endogenous Technological Change', *Journal of Political Economy*, **98**, 71–102.
Saint-Paul, G. (1997), 'Business Cycles and Long-run Growth', *CEPR Discussion Paper*, 1642.
Stadler, G.W. (1990), 'Business Cycles Models with Endogenous Technology', *American Economic Review*, **80**, 763–78.
Stiglitz, J. (1993), 'Endogenous Growth and Cycles', *NBER Working Paper*, 4286.
Thirlwall, A.P. (1979), 'The Balance of Payments Constraint as an Explanation of Inter-

national Growth Rates Differences', *Banca Nazionale del Lavoro Quarterly Review*, **32**, 45–53.
Tong, H. (1983), *Threshold Models in Non-linear Time Series Analysis*, Berlin: Springer-Verlag.
Young, A. (1928), 'Increasing Returns and Economic Progress', *Economic Journal*, **38**, 527–42.

6. Modelling historical growth: a contribution to the debate[1]

Mark Roberts

6.1 INTRODUCTION

In the last decade, a debate has emerged concerning both the extent to which Nicholas Kaldor's writings on cumulative causation (CC) are consistent with his well-known critique of 'equilibrium economics' (see, *inter alia*, Kaldor, 1972, 1975, 1977) and the extent to which 'formal' (i.e. mathematical) modelling is able to capture Kaldor's resulting vision of growth as a historical process. In particular, whilst Kaldor himself believed CC to provide for a theory of historical growth in keeping with what he called for in his criticisms of 'equilibrium economics', Setterfield (1997a; see also 1997b, 2002, p. 230) has argued that this is not so because such causation represents a mechanical process incorporating only positive feedback, with the result that, at best, history only matters insofar as initial conditions matter.[2] Consequently, Setterfield has argued that Kaldor's writings display a tension, a tension best illustrated by the fact that formalisations of his CC schema possess a determinant equilibrium solution as defined by Setterfield (1997b, p. 55).[3] That is to say, a solution that is both uniquely defined in terms of exogenous data and that is reached independently of the path taken towards it. To overcome this tension, Setterfield sought to extend one such formalisation of Kaldor's CC framework, namely the 'standard' CC model of Dixon and Thirlwall (1975), to incorporate a source of negative feedback, judging the result to be a 'more generally hysteretic' model.[4]

However, Setterfield has, in turn, been criticised on two grounds by Argyrous (2001) and Toner (2001). First, Toner claims that there is, in fact, no tension between Kaldor's writings on CC and his critique of 'equilibrium economics'. Rather, he alleges that the tension is a misperception of Setterfield's arising from his incorrectly taking such 'formal' models as the Dixon–Thirlwall (DT) model as representative of Kaldor's writings on CC. Second, Argyrous argues that, whilst Setterfield is correct that the DT model is ahistoric, his extension is unable to sustain his claim of being a 'more generally hysteretic' model. This being the case, the extended model does not even resolve the tension between 'formal' models of CC and Kaldor's historical growth vision. Indeed, both Argyrous and

Toner suggest that it is impossible to 'formally' model growth in a manner consistent with Kaldor's historical vision. In response, Setterfield (2001, p. 109) has accepted that there are shortcomings with his extended model, but has disputed both the argument that there is no tension in Kaldor's own writings and the suggestion that it is impossible to 'formally' model historical growth.

In this chapter we agree with Setterfield that there is a tension in Kaldor's writings on the irrelevance of 'equilibrium economics' and CC. Given this, the major contribution is then to modify Setterfield's extended CC model to show that it is possible to construct a 'formal' model that is consistent with the *spirit* of Kaldor's historical growth vision. As will be seen, the key to achieving this lies in the adoption of an approach to modelling whose methodology is consistent with the notion that the economy is a system whose structure is open to endogenous transformation. In particular, it can be achieved through the adoption of an open systems-*ceteris paribus* (OS-CP) approach to modelling that Setterfield himself has elsewhere advocated (Setterfield, 1997c, 2003).[5]

The structure of the remainder of this chapter is as follows. In sections 6.2 to 6.4, we revisit the debate between Setterfield and his critics. More specifically, in Section 6.2 we demonstrate the ahistorical nature of growth in the DT model and agree with Setterfield that the resulting tension in Kaldor's writings is not attributable to the failure of modellers to correctly represent his views on the subject. In Section 6.3 we present a diagrammatic version of Setterfield's attempt to 'formally' extend the DT model in a manner that renders it consistent with Kaldor's critique of 'equilibrium economics' and, in Section 6.4, we discuss the limitations of the resulting framework as a historical model of the growth process. Finally, Section 6.5 resituates Setterfield's extended model within the context of the OS-CP approach to modelling and presents our modification of the model before Section 6.6 concludes.

6.2 THE AHISTORICAL NATURE OF CUMULATIVE GROWTH PROCESSES

As indicated, the 'formal' model of CC that Setterfield takes as representative of Kaldor's writings on the subject is the DT model (Dixon and Thirlwall, 1975).[6] As is well known, the model considers a small economy in which growth is export-led. It consists of four basic relationships:

$$y_{i,t} = \gamma_i x_{i,t} \qquad [1]$$

$$x_{i,t} = -\eta_i \pi_{i,t-1} + \delta_i \pi_c + \varepsilon_i y_c \qquad [2]$$

$$\pi_{i,t} = w_i + \tau_i - r_{i,t} \qquad [3]$$

$$r_{i,t} = r_{i,e} + \lambda_i y_{i,t} \quad [4]$$

Equation [1] states that the rate of real output growth in the home economy (y) is a positive linear function of its real rate of export growth (x). Meanwhile, equation [2] relates x to, first, the relative price competitiveness of the home economy's exports which is decreasing in the rate of price inflation of home exports (π) and increasing in the rate of price inflation of exports produced by economies that compete with the home economy on third markets (π_c), and, second, a measure of the rate of real income growth in its main export markets (y_c). Equation [3] then postulates that π is determined via the practice of applying a mark-up on unit labour costs, thus implying that π is increasing with both the rate of nominal wage inflation (w) and the rate of mark-up growth (τ), but decreasing with the rate of labour productivity growth (r). Finally, equation [4] is Verdoorn's law, which specifies r as a separably additive, positive linear function of exogenous labour productivity growth (r_e) and of real output growth. It is, of course, the positive feedback that this equation provides for that makes growth in the model cumulative. As for the parameters, γ denotes the elasticity of y with respect to x, η the own price elasticity of demand for exports, δ the cross price elasticity of demand for exports, ε the income elasticity of demand for exports and λ the Verdoorn coefficient.

To demonstrate the tension that exists between this formalisation of Kaldor's CC schema and his critique of 'equilibrium economics', we can substitute and rearrange equations [1]–[4] to give:

$$r_{i,t} = \{r_{i,e} + \gamma_i\lambda_i[\delta_i\pi_c - \eta_i(w_i + \tau_i) + \varepsilon_i y_c]\} + \gamma_i\eta_i\lambda_i(r_{i,t-1}) \quad [5]$$

which we may portray diagrammatically in the form of a 'cobweb' diagram. Thus, in Figure 6.1, the difference equation is represented by the function F. As can be seen, in graphing the function, we have assumed that $0 < \gamma\eta\lambda < 1$. Given this, it should be clear that the model possesses a determinate equilibrium solution that leaves no role for history. Thus, if the economy's initial labour productivity growth rate is r_0 then it experiences a positive cumulative process, whilst, if its initial growth rate is r_0' then it undergoes a negative cumulative process.[7] However, regardless of whether the cumulative process is positive or negative, the economy ends-up settling down at the same equilibrium growth rate r^*.[8] Clearly, therefore, history has no role to play. The economy's long-run growth rate is independent of both its initial growth rate and the subsequent path that it follows – ultimately, therefore, growth is path independent.

From this, it is obvious why Setterfield (1997a; see also 1997b, 2002, p. 230) argues that there is a tension between Kaldor's vision of growth as a historical process and his belief that CC provides a vehicle for capturing this vision. However, as a caveat, it must be noted that it is not, in fact, clear that Kaldor

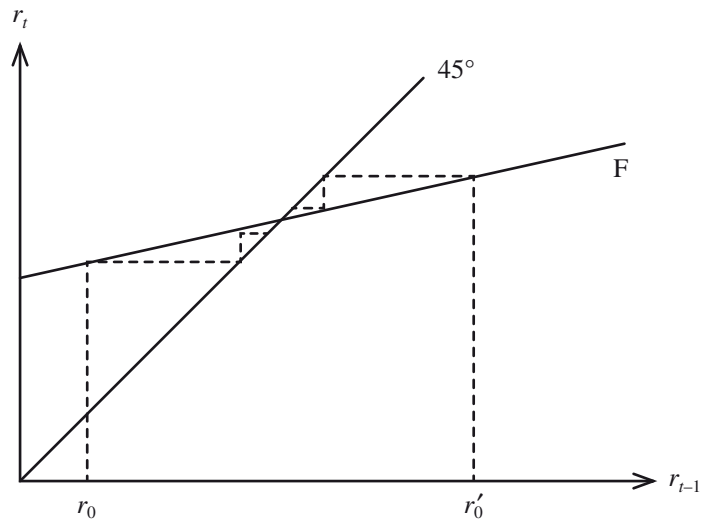


Figure 6.1 Determinate equilibrium solution in the DT model

associated CC with stable dynamics. Rather, there are good reasons to believe that he instead associated it with *un*stable dynamics. Thus, at various points in his pre-1966 work on growth and distribution, a time when he himself was engaged in 'formal' modelling, Kaldor uses the word 'cumulative' to refer to unstable dynamics (see, for example, Kaldor, 1955–1956, pp. 221–222, 1961, p. 249). Furthermore, in a footnote to their original article, Dixon and Thirlwall note that, in private correspondence, Kaldor indicated to them that he believed the case of unstable dynamics to be the one implicit in his own verbal CC models (Dixon and Thirlwall, 1975, p. 208, footnote 2).[9]

Assuming $\gamma\eta\lambda > 0$, the case of unstable dynamics occurs when $\gamma\eta\lambda \geq 1$ and is portrayed in Figure 6.2. In this case, the initial value of *r does* affect the outcome. Thus, if the economy starts-off with a growth rate such as r_0, where $r_0 < r^*$, then failure continually breeds more failure in a negative cumulative process. Conversely, if it starts off with a growth rate such as r_0', where $r_0' > r^*$, then success continually promotes more success in a positive cumulative process. Hence, in the case of unstable dynamics, the DT model does not possess a determinate equilibrium solution and there is a role for history in the form of initial conditions. However, this is the only role for history and to reduce history to initial conditions in this manner is to obviously take a very restrictive view of what constitutes history (Roberts, 2002, p. 87; Setterfield, 1997a, p. 371). More to the point, the treatment of history in this manner does not seem in keeping with the spirit of what Kaldor was calling for in his critique of 'equilibrium economics'.

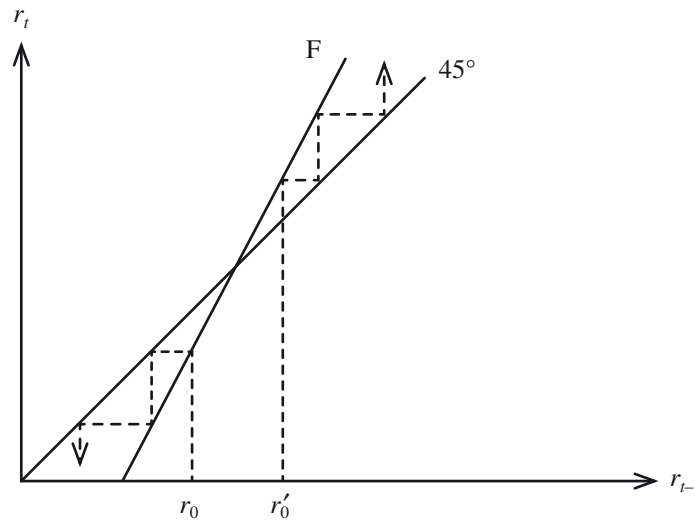


Figure 6.2 Unstable dynamics in the DT model

Thus, the DT model with unstable dynamics does not constitute 'a sequence analysis in which the course of development is dependent on the path of evolution' (Kaldor, 1975, p. 401).[10]

Furthermore, the DT model with unstable dynamics possesses another obviously unsatisfactory property that makes it inappropriate as a historical model of growth and development. This is that the explosive growth/decline implied by such dynamics is clearly unrealistic (see also Swales, 1983, p. 72).[11] In particular, Setterfield (1997a, p. 371, p. 377, 1997b, p. 68) points out that it is inconsistent with a key stylised fact of the long-run growth process. Namely, that in the course of their development, some initially fast-growing regions, such as Britain, have transformed into relatively slow-growing ones.

Toner (2001) has, however, argued that Setterfield is incorrect to take 'formal' CC models, like the DT model, as representative of Kaldor's own writings on the subject. Thus, he argues that, at various points in his work, Kaldor identified a number of alternative, context-specific, mechanisms by which CC might break down.[12] Given their context-specific nature, Toner argues that there is no inconsistency between Kaldor's historical vision of the growth process and his belief that this vision is captured by the notion of CC. Rather, the tension is an apparent one arising from the failure of such 'formal' models as the DT model to adequately capture the subtlety of Kaldor's verbal arguments concerning the nature of the CC process. However, as Setterfield (2001, pp. 108–9) has argued in reply to Toner, the context-specific mechanisms for endogenous breakdown

that Kaldor identified were not a central element of his verbal theorising on CC. This being the case, the tension that exists between Kaldor's critique of 'equilibrium economics' and his views on how CC proceeds is not an artefact of Setterfield incorrectly taking the DT model as representative of these views.[13]

6.3 SETTERFIELD'S EXTENDED CUMULATIVE GROWTH FRAMEWORK

In this section we set out a diagrammatic version of Setterfield's (1997a, 1997b) algebraic attempt to formally extend the DT model in a manner that renders it consistent with Kaldor's historical growth vision.[14] This will provide the basis both for our examination of the arguments of Setterfield's critics and our modification of the model to allow it to more fully realise the claims that Setterfield originally made for it.

To set up his formal attempt at extension of the DT model, Setterfield argues that CC should not be viewed as occurring within a vacuum, but as occurring within a specific institutional and technological context. Furthermore, the institutional and technological context should be thought of as being 'deeply endogenous' to the CC process, that is to say, as being ultimately affected by the process being modelled. In particular, Setterfield argues that increasing economic development is associated with both increasing institutional and technological interrelatedness.[15] The consequence of this positive relationship between interrelatedness and development is that fast growth will not inevitably and mechanically beget continued fast growth.[16] This is because, by leading to increasing interrelatedness, the development of the economy that results from the fast growth brings about the danger of 'lock-in'. Increasing interrelatedness brings about such a danger because it causes 'either the costs of technological and/or institutional change to become prohibitive' or gives rise 'to coordination failures in the context of decentralised decision-making' (Setterfield, 1997a, p. 372). In particular, Setterfield argues that there is a development threshold. Once this threshold is passed, lock-in occurs. At this point, the reigning institutional and technological framework becomes dysfunctional to the maintenance of fast growth, the result being an endogenous fall in λ and/or ε.[17] The result is an endogenous growth slowdown and what Setterfield originally claimed to be a 'more generally hysteretic' model of growth and development (Setterfield, 1997a, 1997b).

To illustrate the above intuition, we can use a 'cobweb' diagram similar to those in Figures 6.1 and 6.2. In doing so, it will be noted that, in keeping with the previous figures, the analysis is couched in labour productivity terms. This contrasts with Setterfield's presentation of his model, which is couched in in-

come level terms. The choice to reformulate the analysis in labour productivity terms is attributable to the fact that an economy's level of labour productivity is a better indicator of its level of development than its level of income. In particular, if it has a large population, an economy can have a high level of income whilst being relatively underdeveloped. It will further be noted that, for simplicity, we assume that lock-in affects only ε. Anyway, consider Figure 6.3 and assume that the institutional and technological lock-in threshold is marked by a corresponding threshold level of labour productivity, R_{th}. Furthermore, say that the economy's initial labour productivity growth rate is r_0 so that it starts-off experiencing a positive cumulative process, but that it passes R_{th} when its growth rate reaches r_2. This being the case, the determinate equilibrium solution r^* towards which the economy was originally heading becomes irrelevant because the economy never actually reaches it. Rather, upon R_{th} being passed, the schedule F shifts downwards to F' as lock-in causes an endogenous fall in ε. This means that r^* ceases to exist and instead the economy experiences a sudden growth slowdown in which its growth rate drops from r_2 to r_3. From here, the economy experiences a further, more gradual, slowing of its growth rate as it heads towards r^{**}.[18]

From Figure 6.3 it should be clear, therefore, that, in Setterfield's extended model, success no longer breeds more success, but rather, consistent with the historical experiences of such countries as Britain, a reversal of fortunes occurs as a result of forces endogenous to the system.

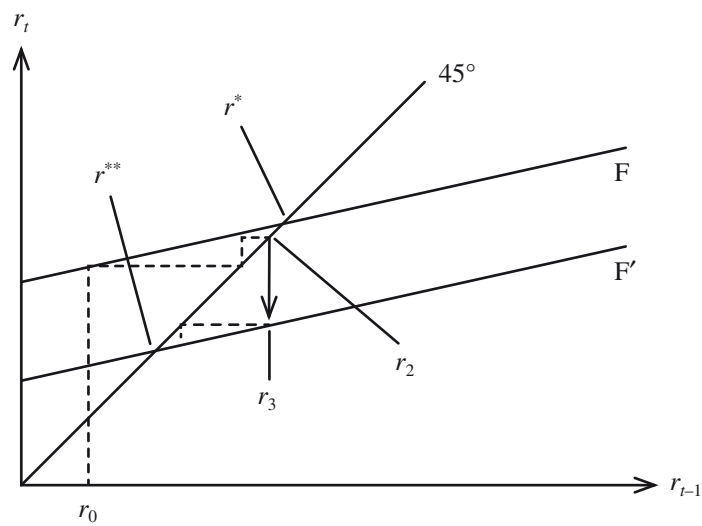


Figure 6.3 Lock-in in Setterfield's extended DT model

6.4 THE LIMITS OF SETTERFIELD'S FRAMEWORK AS A HISTORICAL MODEL

Having set out our alternative presentation of it, we are now in a position to be able to see precisely why Arygrous (2001) in particular has disputed Setterfield's claim that his model is 'more generally hysteretic' than the original DT model. However, before doing so, it is relevant to mention a criticism of Setterfield not made by either Arygrous or Toner (2001). In particular, Setterfield partly based his original claim of having produced a 'more generally hysteretic' model on the fact that, even when $|\gamma\eta\lambda| < 1$, the deterministic equilibrium solution r^* in Figure 6.3 that would normally apply is irrelevant. This being the case, it would seem that history dominates any equilibrium in determining an economy's outcome. However, Figure 6.3, not to mention Figure 6.1, has been drawn based upon empirically realistic values for the parameters of the DT model.[19] However, given these realistic parameter values, we see that the economy starting at r_0 in Figure 6.3 approaches r^* extremely rapidly. Thus, in the first period alone more than one-half of the gap between r_0 and r^* is closed. From this it follows that, unless it starts off very close to R_{th}, an economy is unlikely to pass its lock-in threshold before it reaches r^*. Indeed, if an economy is initially some way below R_{th} then it is likely to remain at r^* for a prolonged period of time before the threshold is passed. This indicates that, *contra* Setterfield, his model does not render the deterministic solution of the DT model completely irrelevant; rather, this solution may be relevant for (a possibly prolonged) period of time.[20]

Moving on to Argyrous' criticism of Setterfield's extended model, he argues that the model does not have a legitimate claim to being a historical model of growth because, ultimately, it still retains a deterministic interpretation. In particular, he interprets Setterfield's model as implying that a growth slowdown must, sooner or later, *inevitably* occur (Argyrous, 2001, p. 405). Furthermore, we may point out that, as presented, the model gives the strong impression that, once it has passed R_{th}, an economy will settle down at a determinate equilibrium solution which represents a strong attractor for the system's dynamics. Thus, it would seem that the economy portrayed in Figure 6.3 ultimately settles down at r^{**}.[21,22]

6.5 A MODIFIED SETTERFIELD MODEL OF HISTORICAL GROWTH

Both Argyrous (2001, p. 105) and Toner (2001, pp. 101–2) question whether the limitations of Setterfield's model in capturing Kaldor's historical vision are not due to a fundamental incompatibility of 'formal' models with such a vision. Setterfield (2001, p. 111), however, replies to this by defending the potential

ability of 'formal' models to capture the Kaldorian vision. Furthermore, Setterfield (2001, p. 109) states that he 'would strongly resist' Argyrous' interpretation of his model as implying that lock-in, and, therefore, a growth slowdown, is inevitable. However, this gives rise to the question: why did Argyrous misinterpret Setterfield in the first place? I would argue that it is attributable to Setterfield's failure to explicitly situate his model within the methodological framework of the open systems-*ceteris paribus* (OS-CP) approach to modelling that he himself has elsewhere advocated (Setterfield, 1997c, 2003). As the first half of its title suggests, this approach to modelling treats the economy as an explicitly open system. By this we mean a system that lacks extrinsic and/or intrinsic closure. A system lacking extrinsic closure is one in which exogenous factors are explicitly allowed to influence system outcomes in a non-defined way, whereas a system lacking intrinsic closure is one in which the same effects do not always inexorably follow from the same causes. The lack of intrinsic closure in an open system follows from the postulated existence of intentional human agency, the potential for human agents to have always acted differently in any given set of circumstances, at the 'deep' level of reality, which is the level of reality at which the tendencies which generate actually observable events exist.

However, as the second half of its title suggests, the OS-CP approach to modelling does not treat all relations as open. Rather, to facilitate modelling, and in line with proper abstraction, the approach introduces closure to some relations by treating them as constant over time, whilst, at the same time, explicitly acknowledging that these relations may be subject to – *a priori* unspecified – endogenous transformation over time. In other words, all closures are explicitly acknowledged as being *conditional* and, therefore, essentially, artificial. Such 'locking up without ignoring' is justified either on empirical grounds ('empirically grounded locking up without ignoring') due to the existence of relatively enduring institutions or on purely analytical grounds ('pure locking up without ignoring') to help in the isolation of the working of the 'deep' tendencies that are of interest.

A corollary of this is that the notion of equilibrium retains a role in the OS-CP approach, but all equilibria are *conditional* equilibria; they are not determinate outcomes, but intermediate positions explicitly acknowledged as being subject to potential revision by forces that are deeply endogenous to the system. Adopting the terminology of the OS-CP approach, it becomes clear that the reason why 'formal' CC models such as the DT model are ahistorical in character, and, therefore, inconsistent with Kaldor's historical growth vision, is that they represent closed systems. That is to say, the key relations, equations [1]–[4], in the DT model are presented as though they are empirical regularities whose continued existence and systematic operation is cast in stone with the result that equilibrium, where it exists, appears as an inevitable and final resting place for

the system. Given this, we seek, in this section, to explicitly resituate Setterfield's model within the methodological framework of the OS-CP approach. Not only this, we extend the model by drawing upon the work of various authors, including Setterfield himself. Thus, we draw both upon the purely verbal suggestions that Setterfield makes in reply to his critics as to why lock-in should not be considered inevitable and why the post-lock-in equilibrium in his model should not be considered a determinate outcome, and upon verbal arguments that he makes elsewhere in his work. This resituation and modification of Setterfield's model within the context of the OS-CP approach serves to provide substance to Setterfield's defence of the potential of producing 'formal' models that are consistent with the spirit of what Kaldor was calling for in his critique of orthodox economics.[23]

To recast Setterfield's extended model within an OS-CP framework we need, first of all, to explicitly point out that, with the exception of ε, we are treating all parameters/equations in the model as invariant over time, despite the fact that we know they are likely to be deeply endogenous to the growth and development process. Thus, for example, we have already specified λ as exogenous despite Setterfield's arguments that it will be subject to endogenous revision if lock-in occurs. Given that we did this purely as an aid to exposition, it represents an example of pure locking up without ignoring. Other examples of pure locking up without ignoring include the modelling of competition between economies as being price based when we know it is more likely to be non-price based (see, most notably, McCombie and Thirlwall, 1994). Meanwhile, our treatment of some parameters/equations as invariant to the growth process also has elements of empirically grounded locking up without ignoring. Hence, the assumption that π is determined via the practice of applying a mark-up on unit labour costs represents the institution of 'normal cost pricing', evidence on the prevalence of which is provided by, *inter alia*, Carlin and Soskice (1990, pp. 140–43). Of course, this practice of determining prices might well be subject to deeply endogenous revision over time, but, given that we are interested in neither the retroduction of the mechanism by which this occurs or its implications for the functioning of capitalism, we abstract from such possibilities.

So much for the application of the *ceteris paribus* part of the OS-CP approach. On the open systems side, we can complete the restituation of Setterfield's model by thinking of R_{th} as being the threshold level of development above which lock-in *tends* to occur rather than as the threshold at which lock-in *inevitably* occurs. In this way, it becomes clear that R_{th} is a relation that lacks intrinsic closure. In particular, intentional human agency keeps open the possibility that lock-in may be avoided through, for example, entrepreneurial ingenuity. By denying intrinsic closure to the lock-in threshold, Setterfield's resituated model obviously also comes to lack overall closure. Furthermore, given that we have explicitly acknowledged above that all relationships are subject to a *ceteris*

paribus clause, it follows that, when the tendency for lock-in is observed at the level of the 'actual', the post lock-in equilibrium takes on the character of a conditional equilibrium that we know is subject to possible future revision by, as yet unspecified, deeply endogenous forces. This overcomes the criticism of Setterfield's extended model that the post lock-in equilibrium represents a determinate, and therefore ahistorical, outcome. Finally, note that the OS-CP approach is perfectly consistent with the fact that the original deterministic solution of the DT model retains some relevance within Setterfield's extended framework. This is, again, because this solution clearly represents only a conditional equilibrium. In fact, the retention of some relevance for the original DT solution actually assists Setterfield's analysis. After all, there have existed historical episodes during which industrial countries have experienced (relative) growth rate stability.

Having recast Setterfield's model within the framework of the OS-CP approach, we can now progress to our further 'formal' modification of the model. In particular, although we now know that the post lock-in equilibrium, where observed, is a conditional equilibrium, the tendencies that might engender its deeply endogenous revision have, as of yet, gone unspecified. At this point, we may start by drawing upon a theory of institutional evolution that has been advocated not only by critical realists (see, for example, Lawson, 1994, p. 520), but also by Setterfield himself (1997b, chapter 5).[24] More precisely, we draw upon the theory that institutions not only provide the framework for intentional action by agents, but are also themselves dependent upon such action for their reproduction. This being the case, institutions are always subject to the possibility of endogenous revision if agents choose to act in a way that does not ensure their unaltered reproduction. More specifically, both Setterfield (1997b, chapter 5) and Cornwall and Cornwall (2001, chapter 6) argue that, at any given point in historical time, the institutions that are in place both constrain and facilitate the actions of agents, so giving rise to economic outcomes that may be judged as either good or bad by agents and their representatives in society. If the judgment is bad then pressure may arise for the modification of existing institutional arrangements thereby creating a negative feedback from intentional action to institutions. In turn, this constrains and facilitates intentional action at future points in historical time and so it goes on.[25]

What the above suggests is modifying Setterfield's model by introducing a growth rate threshold below which there is a *tendency* for agents to become dissatisfied with the prevailing institutional regime; in particular, assume that there exists a *dissatisfaction tendency threshold* of r_{th}. If $r \geq r_{th}$ then all tends to be well – individual workers, shareholders and entrepreneurs, as well as the organisations representing them, tend to be satisfied with the currently prevailing institutional regime.[26] Workers and their representatives tend to be satisfied because real incomes are growing relatively fast and unemployment is relatively

low. Meanwhile, shareholders, entrepreneurs and their representatives tend to be relatively happy because profits are high enough to be judged satisfactory. However, if $r < r_{th}$ then agents tend to become dissatisfied with the prevailing institutional regime because real incomes are growing relatively slowly, unemployment is relatively high and profits relatively low. Consequently, when $r < r_{th}$ pressure arises for reform at both the political level and the level of production itself.[27] The pressure arises at the political level as a result of worker dissatisfaction and at the level of production as a result of dissatisfaction from shareholders and entrepreneurs.

Important to note is that r_{th} is a *metaphor* for a tendency that is assumed to exist at the 'deep' level of reality and which might not actually therefore manifest itself in the guise of successful reform at the 'actual' level of reality. Rather, even abstracting from possible countervailing tendencies at the level of the 'deep', for this to be the case two conditions must be satisfied: (1) sooner or later, the pressure for reform must become effective, and (2) the implemented reforms must be successful. Only if these two conditions are met will structural change occur that shifts the economy into a new institutional regime that is conducive to improved macroeconomic performance.[28,29]

Assuming the above conditions are satisfied, the consequences of introducing the dissatisfaction tendency threshold into Setterfield's 'formal' analysis are illustrated in Figure 6.4. In particular, the consequences are illustrated for an economy whose initial conditions are assumed to be $R < R_{th}$ and $r \geq r_{th}$. Provid-

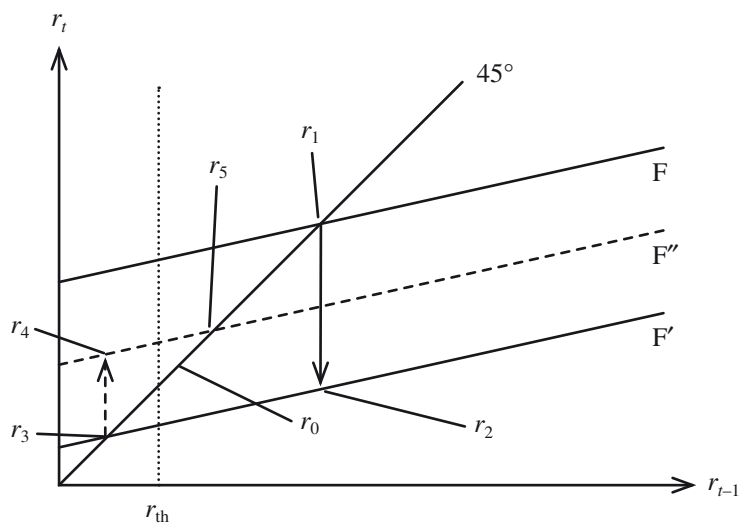


Figure 6.4 A modified Setterfield model of historical growth

ing labour productivity, R, is sufficiently below R_{th}, the economy starts-off with a growth rate of r_0, but quickly converges on the higher growth rate r_1 through a positive cumulative process. Here the economy is in conditional equilibrium and therefore remains at rest, experiencing a period of growth rate stability. This goes on for as long as $R < R_{th}$. However, as soon as $R \geq R_{th}$ the currently prevailing institutional and technological regimes become locked-in. Consequently, they become dysfunctional to the maintenance of fast growth. As in Setterfield's original analysis, this causes the F schedule to shift downwards to F' as ε undergoes an endogenous fall. Given the downward shift in the F schedule, a climacteric occurs as the growth rate first of all suddenly drops to r_2 and then proceeds to tail off further as the economy quickly converges on r_3, which represents a new conditional equilibrium. However, $r_3 < r_{th}$, and so at r_3, pressure for reform exists. Sooner or later, this pressure becomes effective, reform occurs, ε experiences an endogenous increase, and the schedule F' shifts upwards to F''. As a result, the growth rate recovers to r_4 and quickly converges to the conditional equilibrium growth rate of r_5.

However, even confining attention to an economy that starts-off below r_{th}, the sequence of events outlined above arising from the deep tendencies that we have assumed to exist are just one of many possibilities. In particular, if conditions (1) and (2) are not met then the tendencies will result in different effects in the domain of the 'actual'. Thus, for example, consider a situation in which condition (2) is not met. In particular, given a world of fundamental uncertainty, it is conceivable that the reforms implemented in reaction to the condition $r < r_{th}$ becoming binding may have an unanticipated *negative* effect. In this case, they would drive the F' schedule *downwards* rather than upwards to F''. This would cause the growth rate to actually fall further below r_{th}. The result will be a tendency for the pressure for reform to intensify further. *If* those key actors responsible for reform learn from their mistakes (not guaranteed in an open world) then subsequent rounds of reform will presumably be implemented and, eventually, these reforms will succeed in driving the F' schedule sufficiently upwards so as to ensure that $r > r_{th}$. Alternatively, even if conditions (1) and (2) are met, it may be that whilst the pressure for reform does translate into effective reform at the first time of asking, the reforms may not be sufficiently successful to cause r_5 and/or r_4 to lie above r_{th} in Figure 6.4. Again, further rounds of reform would have to follow if the tendency for reform pressure to rise is to be eased.

Above we noted that, once resituated within the OS-CP framework, lock-in should not be regarded as inevitable in Setterfield's extended model because the lock-in threshold is more correctly thought of as being a lock-in *tendency* threshold. In particular, we said that intentional human agency keeps open the possibility that lock-in may be avoided. In some ways, the introduction of the dissatisfaction tendency threshold allows us to 'formalise' this idea. Specifically, consider an economy whose initial condition is $r < r_{th}$ rather than $r \geq r_{th}$. Thus,

for example, an economy for which r_1 in Figure 6.4 lies below r_{th} instead of above it. This being the case, upon $R \geq R_{th}$ coming to hold, the two tendencies modelled would be at work in opposite directions. On the one hand would be the lock-in tendency, exerting downward pressure on the F schedule. Meanwhile, on the other would be the dissatisfaction tendency, working to push the F schedule in the opposite direction. Clearly, in this case, what is observed in the domain of the 'actual' will be open, dependent upon which of these two tendencies dominates.

Clearly, the above modification opens up the number of explicitly possible historical paths in Setterfield's model. However, it is still the case that we should expect the economy to settle down to a conditional equilibrium whose conditionality rests upon exogenous factors. Therefore, from this we may deduce that if the lock-in and dissatisfaction tendencies were the only two tendencies in existence then we would expect an economy to settle down at a steady long-run growth rate. Thus, returning to Figure 6.4, we see that the economy eventually reaches the conditional equilibrium r_5, at which point the internal dynamics of the model cease. However, at the level of the 'actual' it has been observed that capitalistic growth has tended to occur in eras. In particular, for the advanced economies, Cornwall and Cornwall (2001) have identified four distinct eras of macroeconomic performance since the late 19th century – an extended boom from the end of the 19th century to the end of the 1920s, the Great Depression of the 1930s, the Golden Age of Capitalism between 1945 and 1973, and, what Setterfield and Cornwall (2002) refer to as, 'The Age of Decline' from 1973 to the present day.[30] It is possible to further extend Setterfield's analysis to incorporate deep tendencies that cause the model to generate such episodic growth without necessarily claiming that these have been the precise tendencies accounting for the endogenous transformation through the four distinct eras of performance identified. In particular, it is possible to further extend the analysis to generate growth cycles whose amplitude and period are not fixed, but are historical artefacts discernible only *ex post*. More specifically, we can extend the model by introducing two additional assumptions. Our first additional assumption is that the lock-in tendency threshold is not exogenously given, but that, consistent with arguments presented by Setterfield (2001, p. 110), that it is an endogenous variable. Hence, we assume R_{th} evolves endogenously in the manner:

$$(R_{th,\,1},\, R_{th,\,2},\, R_{th,\,3},\ldots,\, R_{th,\,\infty})$$

where

$$R_{th,\,n+1} > R_{th,\,n}$$

where the threshold is indexed according to the number of the current institutional and technological regime in the sequence of regimes that have been historically passed through. That is to say, we assume that the level of labour productivity at which lock-in tends to occur is sensitive to the current institutional and technological regime occupied, and that it increases – by a regime sensitive amount – as the economy moves from regime n to regime $n + 1$.[31]

Meanwhile, our second additional assumption, not inspired by Setterfield, is that unless the economy progresses from regime n within a certain timespan then the conduciveness of the regime to continued growth will display a tendency to deteriorate. We may 'formally' specify this as follows:

$$\text{if } t_n < T_n, \text{ then } \varepsilon_n = \bar{\varepsilon}_n$$

$$\text{if } t_n \geq T_n, \text{ then } \varepsilon_n = f_n(\bar{\varepsilon}_n, t_n - T_n)$$

where

$$\varepsilon_n > 0$$

$$f_n(\bar{\varepsilon}_n, t_n - T_n) = \bar{\varepsilon}_n \text{ when } t_n = T_n$$

$$\frac{\delta f_n(\cdot)}{\delta(t_n - T_n)} < 0$$

Hence, if regime n has been in place for a length of time $t_n < T_n$ then ε remains constant. However, after T_n, the conduciveness of regime n to continued growth displays a tendency to deteriorate as ε starts to diminish according to an *a priori* unspecified functional form specific to the regime. This may be justified by the fact that an economy does not exist in isolation and hence its relative non-price competitiveness is not independent of developments in other economies. Rather, other economies will themselves be evolving through different institutional and technological regimes. This being the case, if the economy fails to progress from regime n then, after a certain time that will be context sensitive, we can expect its relative non-price competitiveness and hence ε to tend to diminish as other economies improve their relative levels of non-price competitiveness by adopting new regimes.[32]

That the incorporation of these two additional tendencies prevents our system from settling down at a conditional equilibrium whose conditionality is dependent upon factors external to the model (but endogenous to the growth and development process more generally) and instead leads to the internal generation of cyclical growth of an unpredictable nature can be demonstrated by reconsidering the economy in Figure 6.4. Without our two additional assumptions, it will

be recalled that following lock-in to its original institutional and technological regime, which we may now label regime $n = 1$, this economy experienced reform that caused it to move into a new regime, $n = 2$. However, with regime $n = 2$ in place, the economy came to settle down at the growth rate r_5, a conditional equilibrium that we acknowledged as being likely to be subject to revision without explicitly specifying the tendencies that might make for such revision. However, with our two additional tendencies, this no longer remains true. Instead, our first additional tendency means that the economy now faces a new lock-in tendency threshold, $R_{th,2}$. If the economy passes this threshold, and the tendency is realised, before $t_2 \geq T_2$ then the schedule F'' will shift downwards.[33] Depending upon the seriousness of the lock-in, this may or may not cause the growth rate to fall sufficiently to take it back below r_{th}. If it does, then reform, or a sequence of reforms, will occur that eventually lead the economy's growth rate to rise once again above r_{th}. However, this result will be dependent both upon whether those responsible for reform learn from any mistakes that they make and upon the reforms coming to fruition before $t_2 \geq T_2$. Furthermore, even if these two conditions are met and the growth rate does rise back above r_{th} as the economy enters regime $n = 3$, any resulting stability will only tend to last for as long as the economy remains below the new lock-in tendency threshold, which, upon entering regime $n = 3$, will have increased to $R_{th,3}$, and for as long as $t_3 < T_3$.

Alternatively, consider the case where, back at r_5 in Figure 6.4, the economy fails to reach $R_{th,2}$ before $t_2 \geq T_2$.[34] This opens up a whole different set of possible paths that the economy may follow. Thus, one possibility is that its deteriorating relative non-price competitiveness may shift the economy below r_{th}, thereby creating pressure for reform that may succeed in moving the economy into regime $n = 3$ before the $R_{th,2}$ threshold has been passed. Other possibilities arise because of the openness of the timing of the economy passing r_{th} relative to it passing $R_{th,2}$, the openness of the pressure for reform in the economy translating into successful reform either at the first attempt or at subsequent attempts and the openness of the timing of any successful reform relative to the economy passing $R_{th,2}$. Furthermore, for each of these possibilities it should be clear that, whichever one results, there are then an equal number of whole new additional possibilities. Hence, ultimately, the internal dynamics of our model are such that there are a continuum of different alternative paths that an economy may follow, and which one it does follow cannot be determined *a priori*, but rather is dependent upon historical context.[35]

6.6 CONCLUSION

From the above it follows that, through resituating Setterfield's model using his own OS-CP approach, we have been able to obtain a model that is perfectly

consistent with the historical vision of the growth process that Kaldor was calling for in his critique of 'equilibrium economics'. In doing so, we have explicitly justified Setterfield's defence against Argyrous and Toner of the possibility of 'formally' modelling open but structured social processes. We have further been able to resolve the tension that exists between the inherently ahistoric nature of CC models and Kaldor's belief that the notion of CC provides a vehicle for capturing his historical vision. Furthermore, with the simple addition of a growth rate threshold below which dissatisfaction and pressure for reform tends to arise, and with the generalisation from one lock-in tendency threshold to a series of regime specific thresholds, we have been able to transform Setterfield's model into one in which conditional equilibrium does not ultimately come to be dependent upon unmodelled factors. Rather, the internal dynamics resulting from the interaction of the postulated deep tendencies result in the expectation that capitalism will exhibit a deep *tendency* towards cyclical growth, with the amplitude and period of each cycle being historical artefacts discernible only *ex post*. Such a deep tendency contrasts with the smooth path of growth that neoclassical growth theory leads us to expect.

NOTES

1. This chapter is based upon chapter 4 of my PhD thesis, of which Tony Thirlwall was the external examiner. I am grateful to both New Hall and the Department of Land Economy, University of Cambridge, for financial support and to John McCombie and Mark Setterfield for comments upon the chapter of my PhD on which this chapter is based.
2. Related to this, Hargreaves Heap (1989, p. 142) has argued that CC represents a model of historical *continuity* rather than historical *change*, whilst Gordon (1991) has criticised CC models for being characterised by 'too much cumulation' (see Setterfield, 2002).
3. In defining a determinate equilibrium, Setterfield draws upon Kaldor (1934).
4. In this chapter we follow Setterfield (1997a, 1997b) by using the word hysteretic to describe a model that allows more of a role for history than simply the importance of initial conditions.
5. Those adopting an overtly critical realist position also frequently argue that it is impossible to 'formally' model open but structured social processes. However, Setterfield (2003) contends that the OS-CP approach is consistent with the ontological and methodological premises of critical realism.
6. This model was developed as an explicit attempt to formalise the mechanism of CC described in Kaldor (1970).
7. It follows from this that, when its dynamics are stable, the DT model does not, contrary to what is often asserted, necessarily predict that success (failure) inexorably breeds more success (failure). Rather, the model only predicts that an economy with a fast initial growth rate will go on to experience even faster growth *if* the fast initial growth rate is low *relative* to the economy's equilibrium growth rate.
8. $r^* = \{r_e + \gamma\lambda[\delta\pi_c - \eta(w + \tau) + \varepsilon y_c]\}/(1 - \gamma\eta\lambda)$, thereby implying that r^* is defined purely in terms of exogenous data.
9. Other authors also associate CC with unstable dynamics (see, for example, Swales, 1983, p. 72).
10. The narrow conception of history as the importance of initial conditions is the same as that which appears in multiple equilibria models or models with 'chaotic' dynamics (see also Setterfield, 1997c, pp 63–65).

11. When $\gamma\eta\lambda = 1$ and the F schedule coincides with the 45-degree line then we have a role for history in the form of initial conditions without the undesirable property of explosive growth or decline. However, this is, obviously, a very special case.
12. Toner (2001, pp. 98–100) gives three examples of such mechanisms. Roberts (2002, p. 89) provides further examples.
13. A possible exception to this is the *emergence* of a binding balance-of-payments constraint as a mechanism for the endogenous breakdown of a positive cumulative process. The DT model has been extended to include a balance-of-payments constraint, but the constraint is always assumed to be binding (see McCombie and Thirlwall, 1994, pp. 429–34; Swales, 1983, pp. 77–9). Consequently, the effects of the *emergence* of such a constraint are not captured.
14. Setterfield actually extends a slightly modified version of the DT model in which y_t is expressed relative to y_c, and in which π_c is expressed as an endogenous function of y_c owing to the practice of mark-up pricing by 'foreign' as well as 'domestic' producers.
15. Setterfield is clearer as to why there should be a positive relationship between technological interrelatedness and the level of development than as to why there should be between institutional interrelatedness and the level of development.
16. This statement implicitly assumes that $r < r^*$ ($r > r^*$) if the dynamics of the DT model are stable (unstable).
17. Both Kaldor (see, for example, Kaldor, 1966, p. 9) and Verdoorn himself (Verdoorn, 1956), have drawn upon the idea of a learning function in attempting to explain Verdoorn's law. In this context, technological lock-in can be interpreted as causing an endogenous fall in λ because the opportunities for learning within the current technological regime have ceased.
18. From this alternative presentation it is clear that Setterfield implicitly assumes that the threshold, R_{th}, at which institutional lock-in occurs is the same as that at which technological lock-in occurs.
19. In particular, the gradient of each F schedule has been constructed assuming the empirically reasonable values of $\gamma = 1$, $\eta = 0.5$ and $\lambda = 0.5$.
20. A caveat to this criticism is that whilst $\gamma = 1$, $\eta = 0.5$ and $\lambda = 0.5$ may be empirically realistic for the current institutional and technological regimes, this does not mean they will continue to be so for future regimes. In particular, Setterfield's claim is that λ will be different in different regimes. Nevertheless, it still follows that the deterministic solution of the DT model will be relevant for what might be a prolonged period of time for both the current regime and for future regimes in which λ is reasonably high.
21. This assumes that, once the growth slowdown has occurred, the dynamics are stable.
22. Argyrous (2001), as well as Toner (2001), also criticises Setterfield for emphasising only the negative aspects of increasing interrelatedness to the neglect of the positive features, in particular, the increased division of labour, that are such a feature of Kaldor's CC writings. However, as Setterfield (2001, pp. 107–8) notes, such positive features were always implicit in his model. This is because he builds upon the DT model, a central feature of which is Verdoorn's law.
23. Setterfield (2002) has himself re-formulated his model using the OS-CP approach. However, he does not undertake the further extensions of his model that we present here.
24. Setterfield refers to the theory as the 'institutional hysteresis' model, whilst in critical realist writings it is referred to as the 'transformational model of social activity'.
25. Economics itself may have a part to play in this process. Thus, poor macroeconomic performance leads economists to concentrate on the reasons for this poor performance, leading to the advocacy of policy measures which, if implemented, can help shift the economy to a new regime.
26. As in Setterfield (1997b, p. 85), an 'institutional regime' may be defined as 'a collectivity or matrix of institutions, which forms the current socio-economic environment within which individual agents act, and which helps define relations between these agents'.
27. In keeping with the OS-CP approach, r_{th} must be explicitly acknowledged as being open to exogenous influence. For one, we can expect a tendency for r_{th} to be increasing with the 'world' growth rate because one way in which agents are likely to judge the growth performance of their economy is by comparing it to that of other economies. Comparisons with other economies at similar levels of development are likely to be particularly relevant in this context.

Another factor that may influence r_{th} is the nature of government; i.e. whether rule is democratic or not. Given that we can expect these exogenous factors to impinge upon r_{th} in a non-definable way, the threshold lacks extrinsic closure.

28. It seems reasonable to argue that if $r < r_{th}$ then not only will agents tend to become dissatisfied with the prevailing institutional regime, but also with the prevailing technological regime. Therefore, we can expect a tendency for reform not only to be aimed at the institutional regime, but also at the technological regime. This is part of the reason why we later simplify by assuming strict complementarity between institutional and technological regimes.
29. Both the dissatisfaction tendency and lock-in tendency thresholds can be related to Kaldor's concept of the 'technical dynamism' of an economy. By 'technical dynamism' Kaldor meant 'both inventiveness and *readiness to change or to experiment*' (Kaldor, 1961, p. 265; emphasis added). This is because lock-in implies an absence of readiness to change and experiment, whilst reform, by definition, necessitates a re-emergence of such readiness. Also related is Kaldor's view of ε as being dependent on '… the innovative ability *and adaptive capacity* of its [a country's] manufacturers …' (Kaldor, 1981, p. 340; emphasis added).
30. Setterfield and Cornwall (2002) also consider the possibility that the 1990s witnessed the emergence of a new, neoliberal, growth era, at least for the United States.
31. In a case of pure locking up without ignoring, we assume from now on that institutional and technological regimes are strictly complementary. This makes it valid to refer to a single institutional *and* technological regime.
32. It may be argued that already implicit in the original Setterfield model is the assumption that the timing of the passing of R_{th} coincides with the emergence of new institutional and technological regimes in competitor economies. Otherwise, upon R_{th} being passed there would seem to be little reason for ε to undergo an endogenous fall. This being so, our relaxation of this assumption means that lock-in must be causing a slowdown solely through λ for the reasons given in note 17.
33. We can, approximately, say that this will be the case if $[\ln(R_{th,2}) - \ln(R_5)]/r_5 < T_2$, where R_5 is the level of labour productivity that prevails in the economy at the point in time when it first reaches r_5.
34. Approximately, this will be the case if $[\ln(R_{th,2}) - \ln(R_5)]/r_5 \geq T_2$.
35. For simplicity, our analysis assumes that the threshold r_{th} is determined exogenously of the institutional and technological regime occupied. Again, this is an example of pure locking up without ignoring.

REFERENCES

Argyrous, G. (2001), 'Setterfield on Cumulative Causation and Interrelatedness: A Comment', *Cambridge Journal of Economics*, **25**, 103–106.

Carlin, W. and Soskice, D. (1990), *Macroeconomics and the Wage Bargain: A Modern Approach to Employment, Inflation and the Exchange Rate*, Oxford: Oxford University Press.

Cornwall, J. and Cornwall, W. (2001), *Capitalist Development in the 20th Century: An Evolutionary-Keynesian Analysis*, Cambridge: Cambridge University Press.

Dixon, R. and Thirlwall, A.P. (1975), 'A Model of Regional Growth Rate Differences on Kaldorian Lines', *Oxford Economic Papers*, **27**, 201–14.

Gordon, D. (1991), 'Kaldor's Macro System: Too Much Cumulation, Too Few Contradictions', in Nell, E.J. and Semmeler, W. (eds), *Nicolas Kaldor and Mainstream Economics: Confrontation or Convergence?*, New York: St Martin's Press.

Hargreaves Heap, S.P. (1989), *Rationality in Economics*, Oxford: Basil Blackwell.

Kaldor, N. (1934), 'A Classificatory Note on the Determinateness of Equilibrium', *Review of Economic Studies*, **1**, 122–36.

Kaldor, N. (1955–56), 'Alternative Theories of Distribution', reprinted in Targetti, F. and Thirlwall, A.P. (eds), *The Essential Kaldor*, New York: Holmes and Meier.
Kaldor, N. (1961), 'Capital Accumulation and Economic Growth', reprinted in Targetti, F. and Thirlwall, A.P. (eds), *The Essential Kaldor*, New York: Holmes and Meier.
Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the United Kingdom*, Cambridge: Cambridge University Press.
Kaldor, N. (1970), 'The Case for Regional Policies', *Scottish Journal of Political Economy*, **17**, 337–48.
Kaldor, N. (1972), 'The Irrelevance of Equilibrium Economics', *Economic Journal*, **82**, 1237–55.
Kaldor, N. (1975), 'What is Wrong with Economic Theory', reprinted in Targetti, F. and Thirlwall, A.P. (eds), *The Essential Kaldor*, New York: Holmes and Meier.
Kaldor, N. (1977), 'Equilibrium Theory and Growth Theory', reprinted in Targetti, F. and Thirlwall, A.P. (eds), *The Essential Kaldor*, New York: Holmes and Meier.
Kaldor, N. (1981), 'The Role of Increasing Returns, Technical Progress and Cumulative Causation in the Theory of International Trade and Economic Growth', reprinted in Targetti, F. and Thirlwall, A.P. (eds), *The Essential Kaldor*, New York: Holmes and Meier.
Lawson, T. (1994), *Economics and Reality*, London: Routledge.
McCombie, J.S.L. and Thirlwall, A.P. (1994), *Economic Growth and the Balance-of-Payments Constraint*, London: Macmillan.
Roberts, M. (2002), *Essays in Cumulative Causation*, unpublished PhD thesis, Department of Land Economy, University of Cambridge.
Setterfield, M. (1997a), '"History versus Equilibrium" and the Theory of Economic Growth', *Cambridge Journal of Economics*, **21**, 365–78.
Setterfield, M. (1997b), *Rapid Growth and Relative Decline: Modelling Macroeconomic Dynamics with Hysteresis*, London: Macmillan.
Setterfield, M. (1997c), 'Should Economists Dispense with the Notion of Equilibrium?', *Journal of Post Keynesian Economics*, **20**, 47–76.
Setterfield, M. (2001), 'Cumulative Causation, Interrelatedness and the Theory of Economic Growth: a Reply to Argyrous and Toner', *Cambridge Journal of Economics*, **25**, 107–12.
Setterfield, M. (2002), 'A Model of Kaldorian Traverse: Cumulative Causation, Structural Change and Evolutionary Hysteresis', in Setterfield, M. (ed.), *The Economics of Demand-led Growth: Challenging the Supply-side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Setterfield, M. (2003), 'Critical Realism and Formal Modelling: Incompatible Bedfellows?', in Downward, P. (ed.), *Applied Economics and the Critical Realist Critique*, London: Routledge.
Setterfield, M. and Cornwall, J. (2002), 'A Neo-Kaldorian Perspective on the Rise and Decline of the Golden Age', in Setterfield, M. (ed.), *The Economics of Demand-led Growth: Challenging the Supply-side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Swales, J.K. (1983), 'A Kaldorian Model of Cumulative Causation: Regional Growth with Induced Technical Change', in Gillespie, A. (ed.), *Technological Change and Regional Development*, London Papers in Regional Science, 12.
Toner, P. (2001), '"History versus Equilibrium" and the Theory of Economic Growth by Mark Setterfield: a Comment', *Cambridge Journal of Economics*, **25**, 97–102.
Verdoorn, P.J. (1956), 'Complementarity and Long-range Projections', *Econometrica*, **24**, 429–50.

7. Endogenous growth theory: a partial critique

Philip Arestis and Malcolm Sawyer

THE CONTRIBUTION OF TONY THIRLWALL TO GROWTH THEORY

Tony Thirlwall has been a consistent critic of endogenous growth theory. His writings on growth theory with his emphasis on the role of demand, balance of payments constraints, endogeneity of the natural rate and the role of sectoral analysis have made major contributions to understanding the growth process, and they also provide a continuing (sometimes implicit, sometimes explicit) critique of endogenous growth theory. It is, thus, very fitting to indulge in a critique of endogenous growth theory in this Festschrift. The critique is partial in both the sense of not being comprehensive which is precluded by space considerations and in the sense that we are partial to the work of Tony!

We begin with a brief critical review of endogenous growth theory (EGT). The next section focuses on the assumptions made in EGT on returns to scale and the implications of those assumptions. This is followed by a discussion on the approach to investment and productivity growth in EGT. Section 4 contains a brief review of empirical evidence, followed by a concluding section.

A BRIEF REVIEW OF ENDOGENOUS GROWTH THEORY

It is generally recognised that a revival of interest in growth theory within the mainstream literature was stimulated by papers such as Romer (1986) and Barro (1991), which launched what is now usually termed as Endogenous Growth Theory (hereafter EGT). Other labels were also given (e.g. post-classical growth theory) but the term 'endogenous' was particularly significant as this approach appeared to have an analysis of an endogenously determined rate of growth, rather than the effectively exogenously determined 'natural' rate of the Solow-Swann neo-classical model (Solow, 1956; Swann, 1956). The term 'new growth theory' is also used, but as many have argued (e.g. Fine, 2000; Thirlwall, 2003) most of its ideas have a long history.

Endogenous growth theory has spawned a very large literature, and the issue arises as to what constitutes EGT. The key features of EGT include the following:

1. Capital is to be viewed in a broad sense to include human capital, social capital, intellectual capital, public infrastructure as well as physical capital (whereas the neo-classical approach was largely limited to physical capital).
2. There are constant returns to capital viewed in this broad manner (or at least there is a distinct lower bound to the marginal product of capital, whereas in the neo-classical approach marginal product of capital would tend to zero as the capital-to-labour ratio rose), and overall increasing returns to capital and labour combined.
3. Some capital is privately owned and the returns on this capital accrue to its owners. But there are returns to capital which do not accrue to individual owners: this may be because of externalities arising from privately owned capital, because some capital has no owners (e.g. social capital) or because some capital is owned by the government.
4. Savings arise from (lifetime) optimising decisions by households and lead to an equal amount of investment, with savings automatically flowing into capital formation.
5. Investment in privately owned capital is undertaken for profit-maximising reasons and responds to relative prices, specifically the rate of interest. In effect, savings and investment are brought into equality by the rate of interest (and specifically not through variations in the level of economic activity which is presumed to be at the full-employment level).

The simplest representation of the EGT is given by the so-called AK model (e.g. see Barro and Sala-i-Martin, 1995), which can be used to illustrate some of the main features of the EGT. In this AK model, $Y = AK$ where Y is real output, and K is a measure of capital that remains to be precisely defined, and investment equals savings, i.e. $dK/dt = sY$, where s is the propensity to save. Hence $(1/K)(dK/dt) = sA$ and also $(1/Y)(dY/dt) = sA$ – which is comparable to the warranted growth rate of Harrod, since $A = 1/\nu$. The growth of labour force is taken as n (often the growth of the labour force is taken equal to the rate of population growth). By deduction, the growth of labour productivity, $g_p = sA - n$, which means that provided $sA > n$ the growth of productivity can continue in perpetuity.

The AK model 'does not require the assumption of full employment of labor: since labor is not productive, the level of employment of labor does not matter; indeed some version of the model can omit labor altogether' (Dutt, 2003, p. 73–4) But he later notes that

> other formulations [that] allow for new products […] introduce imperfect competition, thereby allowing them to have neoclassical optimization microfoundations with increasing returns even without externalities, and also derive (at least temporary) returns from innovations that subsequently become public goods. It should be notes that all of these version of new growth theory assume explicitly that labor is fully employed: for instance, in models with research and development activity, the total labor force at any moment in time is engaged either in production or in research and development. (pp. 75–6)

In the Harrod–Domar model, the warranted rate of growth, $s/\nu = sA$ in current notation, and the natural rate of growth n were treated as generally unequal, and the warranted growth path was viewed as having 'knife-edge' problems and exhibiting a degree of instability (cf. Hussein and Thirlwall 2000). It is well known that the neo-classical growth model was perceived to have resolved these problems, in part through the adjustment of the capital–output ratio to bring s/ν and n into equality, and by the removal of the accelerator-type investment equation, which had been at the source of the instability elements of the Harrod–Domar model.

It is well known that the Harrod–Domar approach incorporates an investment function of the simple accelerator type. The equality of savings and investment in the context of that model provided the warranted rate of growth equal to s/ν. The Harrod–Domar approach threw up the instability/knife edge problem in that the warranted growth path appeared to be unstable, at least in the simple formulations. Growth slightly faster than the warranted rate of growth would push up investment, leading to faster growth, and so on. In the context of growth theory this instability issue has never been fully resolved.

The other issue of the Harrod–Domar model concerned the lack of equality between the warranted rate of growth (s/ν) and the natural rate of growth (n): three coefficients s, ν, n separately determined and there appeared no reason why s/ν should be equal to n. The neo-classical solution of Solow (1956) and Swann (1956) was to postulate that ν (the capital–output) ratio would adjust, with substitution between labour and capital depending on relative prices. It was argued that this approach assumes that all savings will be invested (in terms of planned savings and investment) such that $s = n\nu$. Further, the investment function of the accelerator form is withdrawn and there is no independent investment function. In the EGT approach, there is a demand for capital as a function of relative prices (from marginal product of capital being equal to the rate of interest, and marginal product of labour being equal to the real wage). The investment function adjusts capital stock towards the desired level of capital stock, albeit that the speed of adjustment depends on the amount of savings being undertaken. In the EGT approach this issue of instability is sidestepped, as there is no accelerator type mechanism (whether of the form of changes in output or the level of capacity utilisation influencing investment). In general, savings and

investment are brought into equality (at a constant rate of employment[1]) through variations in the rate of interest. In this simple model, investment (as the rate of change of the capital stock) would be linked with the rate of change of output, but the capital stock is assumed to smoothly expand and in the simple presentation there is nothing that could be labelled an independent investment function. The equality between the warranted rate and the natural rate is assured in this simple EGT approach by adjustment of labour productivity, since from above $sA = n + g_p$. Since output and the capital stock are growing faster than the labour force, the capital–labour ratio is continually rising. This can continue in part because the rise in the capital–labour ratio is presumed (by construction) to be generating rising labour productivity. Further, in this simple AK model, with a marginal productivity assumption, all income is paid out as profits. Hence the rate of profit is equal to $Y/K = A$, and this remains constant over time.

One feature of the AK model arises from applying marginal productivity considerations, namely that marginal product of capital equals average product, or $dK/dY = A = Y/K$, and hence payments to capital owners based on marginal productivity would exhaust the total product, leaving wages equal to zero. This may be rescued to some degree by thinking of K as including human capital, such that some payments accruing to capital go to workers as the 'owners' of human capital. But this raises a more substantial issue in the EGT literature, namely that since increasing returns are assumed, the determination of the distribution of income generates a problem since it is well known that payments according to marginal product would more than exhaust the total product in the presence of increasing returns to scale. Payment according to marginal product would more than exhaust product, so we are left to believe that 'owners'/suppliers of human capital are not rewarded at all. In the Cobb–Douglas production function (and others), then, there is a positive marginal product of labour and also of capital, but the private returns to capital are lower than the social returns. In this case (under the constant returns to the private factors assumption) it is the spillover of capital which receives no reward.

The savings propensity can in a sense be readily endogenised so as to arise from household inter-temporal utility maximisation; undertaking such an exercise is significant in three ways:

(i) it portrays the individual as a 'ruthless' optimiser who has sufficient knowledge of the future course of the economy to be able to undertake optimising calculations;
(ii) based on household utility, different paths of the economy can be compared in economic welfare terms, and specifically a 'competitive' market outcome compared with a 'planning' outcome;
(iii) it emphasises that savings are viewed as undertaken by households in pursuit of their own interests, though since firms are owned by households

the view can be taken that firms act in the interests of their owners when undertaken savings (i.e. making retained earnings).

Households are assumed to optimise an inter-temporal utility function, which Barro and Sala-i-Martin (1995) take as having the specific form as:

$$U = \int_0^{\infty} e^{-(\rho - n)t} \left[\frac{c^{(1-\theta)} - 1}{1 - \theta} \right] dt \text{ subject to } \dot{a} = (r - n)a + w + c$$

where c is consumption per capita, n is growth of population, r rate of interest, a assets per person and ρ is discount rate. In the closed economy case, the growth of consumption from this optimisation is given by

$$\frac{\dot{c}}{c} = \frac{1}{\theta}(r - \rho)$$

This approaches faces a range of difficulties. Solow has written that 'I find that I resist this practice instinctively. It seems to me foolish to interpret as a descriptive theory what my generation learned from Frank Ramsey to treat as a normative theory, a story about what an omniscient, omnipotent, and nevertheless virtuous planner would do' (Solow, 1997, p. 12).[2] Dutt (2003) states that 'one can argue that at best what these models [based on infinitely lived consumers maximising discounted utility level under the assumption of perfect foresight] do is to allow a comparison of the actual outcome for economies with some social optimum. But even here its value is limited by the assumption that preferences are given, whereas during the growth process one can expect preference to changes, arguably in unknowable ways' (Dutt, 2003, p. 76). The comparison with a centrally planned economy is of doubtful use when it is considered that 'a centrally planned system must be expected to generate quite different patterns of knowledge to those arising in a decentralized market economy' (Metcalfe, 2001, p. 571). It can be further added that as one of the features of growth clearly is that new products and processes are developed, the meaning of 'constant preferences' is far from clear when the product space over which the preferences are expressed is continually changing. This approach to savings also assumes that it is households which are 'in charge' of the savings process, and that corporations are merely the agents of households so that any savings which they undertake are made on behalf of households. Any notion that savings by corporations are driven by their appetite for funds for investment is clearly absent.

On the production side, the typical firm can be assumed to operate according to a production function of the form:

$$Y_i = F(K_i, A_i L_i)$$

with constant returns to scale at the firm level (cf. Barro and Sala-i-Martin, 1995, pp. 146–7). In the case where there are spillover effects from one firm to another, then A_i is proxied by K and hence:

$$Y_i = F(K_i, KL_i)$$

'In equilibrium, all firms make the same choices' (p. 148), which enables the private marginal productivity (eqation 4.27 in Barro and Sala-i-Martin, op. cit.) to be derived which is given by:

$$f_i(k_i, K) = f^*(L) - L.f^{*\prime}(L)$$

where f is the production function per unit of labour, f_1 denotes first derivative with respect to the first argument, f^* is the average product of capital, and a dash denotes first derivative of the function. Net (of depreciation) marginal product of capital is equated with rate of interest, hence:

$$r = f_1(k_i, K) - \delta$$

Combining these elements, the growth of consumption is then given by:

$$\frac{\dot{c}}{c} = \frac{1}{\theta}[f^*(L) - Lf^{*\prime}(L) - \delta - \rho]$$

and the proportionate rate of change of $k = K/L$ is given by:

$$\frac{\dot{k}}{k} = f^*(L).k - c - \delta k$$

which in effect indicates that the growth rate of k is equal to the output per person minus consumption and depreciation per person. In the Cobb–Douglas production function case the growth of consumption becomes

$$\frac{\dot{c}}{c} = \frac{1}{\theta}[A\alpha L^{1-\alpha} - \delta - \rho]$$

This is a potentially significant result in that the scale (as reflected in the size of the labour force L) of an economy influences the trend rate of growth (of the capital stock). Thus a larger economy benefits not only in terms of static economies of scale but also in growth terms. This can be readily illustrated in the case

of an aggregate Cobb–Douglas production function of the form $Y = KL^a$: if L is treated as 'raw' labour (and hence skills and training included as 'human capital' in K) then a value of a around 0.33 is often used, whereas if L is treated as labour skilled and unskilled then a value of a around 0.67 may be appropriate (guided by the share of labour in national income). With this production function, a doubling of scale in terms of factor inputs would lead to a 2^a increase in labour productivity, that is an increase of 25.7 per cent for $a = 0.33$ and 59.1 per cent for $a = 0.67$. Here $f^*(L) = L^a$, and hence a doubling of L would mean this term increasing by 2^a.

The rate of productivity growth g_p is equal to $sA - n$, which suggests a highly significant role for savings in the promotion of growth. This can be illustrated by some rough orders of magnitude. Let $s = 0.16$, $A = 0.25$ (implying capital-output ratio of 4), n =0.01; then $g_p = 0.04 - 0.01$, i.e. 3 per cent – which is plausible. This approach implies if the savings rate were doubled, then the growth rate would double, but at the same time the growth of productivity would more than double; if $s = 0.32$, this would give $g_p = 0.08 - 0.01$, i.e. 7 per cent.

The EGT approach makes heavy use of the representative firm, and in effect the basic nature of the production function is assumed to be the same across all sectors of the economy (generally, of course, conforming to the increasing returns postulates as outlined above). This contrasts with the approach of Kaldor (e.g. Kaldor 1966) and Thirlwall (e.g. Thirlwall 2002, Chapter 3), where there are significant differences between sectors of the economy in terms of returns to scale and the 'openness' to technical change as well as the conditions of demand (in terms of price and income elasticity). The perceived differences between primary, secondary and tertiary sectors of the economy have often been judged as of particular importance. The 'conversion' of knowledge and new ideas into some form of informational capital and the notion that firms are far sighted in their decisions on inputs and outputs clearly overlooks many of the essential ingredients of the development of new ideas. 'As scholars from Marshall … to Kuznets … have recognized, economic activity changes knowledge directly and indirectly and every change in knowledge opens up the conditions for *changes in activity* and *thus further changes in knowledge, ad infinitum*, and in quite unpredictable ways' (Metcalfe, 2001, p. 570).

THE TECHNOLOGY ASSUMPTION ON RETURNS TO SCALE

The question in this context is the extent of economies of scale, and in particular whether at the firm level the assumption of constant returns to scale can be maintained. It is well known that the presence of economies of scale at the level of the firm undermines any assumption of perfect competition – both in terms

of the stability of the perfectly competitive structure under economies of scale and the 'adding up' problem arising from payment of factors of production according to marginal productivity. The EGT approach generally assumes constant returns at firm level with respect to the private factors of production. The estimates of economies to scale are made at the industry or economy level, whereas the discussion of economies of scale usually involves the enterprise, or lower, level. For example, economies of scale in the car industry on the whole relate to the factory level or to the spreading of design and development overheads at the level of the product, engine say.

The question is raised as to whether the EGT approach depends on such a combination of overall increasing returns to scale and precise constant returns to capital, and there have been different responses. Solow (1994) argues in the following terms, 'Now I return to the question of constant returns to capital. It may not be generally recognized how restrictive this assumption is. There is no tolerance for deviation. Lucas emphasized in his 1988 article that a touch of diminishing returns (human capital in his case) would change the character of the model drastically, making it incapable of generating permanent growth. He did not notice that a touch of increasing returns to capital would do the same, but in a quite different way. … Let net investment be the fraction of s of output so that the time path of K is determined by $dK/dt = s\,f(K)$. It is obvious on the face that there is potential for fairly explosive behavior if $f(K)$ increases more and more rapidly with K. For instance, if $f(K)/K$ increases with K, the rate of growth of K gets faster as K gets larger' (pp. 49–50). Setterfield's (1994) critique is along the lines that suggest 'The treatment of increasing returns to scale in new endogenous growth models is also open to criticism. For example, it is not clear that the existence of increasing returns is a necessary condition for endogenous growth to occur in these models.' (p. 115). Setterfield (op. cit.) then provides an example in which output is function of 'knowledge' with constant returns to 'knowledge'. Dutt (2003, p. 74) gives the example of $Y = AK + BK^{\alpha}L^{1-\alpha}$, which departs from the AK formulation but where the growth of output per head tends asymptotically to $sA - n$.

Barro and Sala-i-Martin (1995) argue that 'the production function can exhibit diminishing or increasing returns to k when k is low, but the marginal product of capital must be bounded from below as k becomes large' (p. 42). Also that 'we showed […] that it was possible to obtain long-term per capita growth without exogenous technological progress if the returns to capital were constant asymptotically' (p. 171). Still, 'we noted in Chapter 1 that the global absence of diminishing returns to capital […] may seem unrealistic, but the idea becomes more plausible if we construe capital, K, broadly to encompass human capital, knowledge, public infrastructure, and so on' (p. 141).

The EGT approach begins at the level of the representative firm, and then aggregates. The question arises as to the appropriate level of aggregation, and

whether the nation state should be treated as the appropriate unit of aggregation. Apart from the convenience that data are available at the level of the nation state, there are a number of arguments for using the nation state as the unit of analysis. A country shares common macroeconomic policies and common social and political institutions and exchange rate, and there is presumed to be substantially greater movement of factors of production and goods and services within a state than between states. In the EGT approach, the emphasis is on the role of externalities, which may spill across national boundaries.

In the case of EGT, because of the effects of scale on both the level of productivity and the rate of growth, finding the appropriate level is a significant issue. For example, suppose we were investigating two areas of comparable size (say regions of a large country, or two neighbouring countries), then if it were appropriate to aggregate them together the predicted level of labour productivity would be 27 per cent higher than if it were appropriate to treat them separately (using the figures from the Cobb–Douglas production function as indicated above). In a similar vein, as also indicated above, there would be a substantial effect on the predicted growth rate. This line of argument would also suggest that there would be major differences between countries in terms of their output per person (for given capital:labour ratio) and their growth rate arising from differences in the country size: compare for example Sweden with the United States with a difference in population of a factor of 30. Similarly the equation above indicating that growth rate depends on the size of the population would point to an acceleration of growth over time as population grows.

If there were, say, N industries, each with economies of scale, then an autarkic economy which sought to operate in all N industries would be subject to economies of scale in the sense that the larger the autarkic economy, the larger would be each industry 'on average'. But in this case the degree of increasing returns is limited by the size of the market, and would be equal to the extent of increasing returns in the 'average' industry. Evidence on returns to scale suggest that whilst many industries operate subject to some increasing returns, others do not (at least above some rather small minimum scale). But in so far as industries do operate subject to increasing returns, these may be of the order of 10 per cent, i.e. a doubling of scale would lead to a 10 per cent reduction in costs or increase in average labour productivity: 'we have found that based on UK industry a *prima facie* case can be made for externalities operating at the level of aggregate manufacturing. No evidence was found for increasing returns which are internal to the industry. To the contrary, at the industry level returns appear to be constant' (Oulton, 1996, p. 111). He reports a range of estimates for externality indexes; the flavour can be given by the following: 'The estimates of α_2 [the externality indicator] range in size from 0.0991 (gross input) to 0.2408 (value added). In other words, an expansion of aggregate manufacturing input by 1% is predicted to raise output by between about 0.10% and 0.24%, holding industry

input constant' (p. 106). Basu and Fernald conclude: 'On both empirical and theoretical grounds, heterogeneity in production appears important for macroeconomics. Although estimates of returns to scale vary widely across relatively disaggregated industries, the average industry appears to produce with constant or even decreasing returns.' (Basu and Fernald, 1997, p. 275).

Unless there are substantial managerial diseconomies of scale, firms can expand through duplication (from the production side), and in that way face something approaching constant returns to scale. From that argument and the evidence just cited, it would seem reasonable to take constant returns to scale with respect to the private factors of production as a starting point, recognising that there would be differences between industries and the minimum efficient scale may be substantial in some industries. This would then indicate that under EGT the externality effects have to be assumed to be substantial, turning constant returns to scale into substantial increasing returns to scale.

One conclusion, which comes from the increasing returns assumption, is that output per person is a function of scale, in that larger economies would, *ceteris paribus*, have higher levels of productivity. This arises from the nature of increasing returns which is postulated, that is having built up from the representative firm, the economy as a whole is subject to increasing returns. Furthermore, there is in effect no demand constraint. In some contrast, a more usual approach would be to postulate that increasing returns operate at the level of the plant or firm, and then up to a limit. The firm's ability to exploit economies of scale may be limited by the size of the market (and Chamberlin's theory of monopolistic competition is a notable example of this).

Barro and Sala-i-Martin (1995) suggest that 'if we can identify L with the aggregate labor force of a country, then the prediction is that countries with more workers tend to grow faster in per capita terms. The empirical results discussed in Chapter 12 for a large number of countries in the post-World War II period indicate that the growth rate of per capita GDP bears a weak positive relation to the size of the working-age population. (These results apply when the initial level of per capita GDP, the average person's education, and some other variables are held constant). Thus, these findings do not reject minor [*sic*] scale effect' (p. 151). The same authors go on to say that the estimated coefficient is not statistically significant (p. 442); also, that they 'have already observed in Chapter 4 that scale effects are not supported empirically if we identify scale with the size of a country's population or economic activity' (p. 220).

The production function approach used in EGT typically involves measures of capital where capital is widely defined to include physical, human, knowledge and social capital. There are then a variety of conceptual and measurement questions which arise, and here we highlight three of them. The first concerns the conceptualisation of human and social capital. We make two brief points here. The first is that human and social capital have proved to be rather nebulous

concepts; for example what constitutes 'human capital': education, skills, marketable skills etc. The second is the notion that social institutions and human knowledge and experience can be treated on a par with physical capital.

The second is a reminder of the still unresolved problems which were brought forward in the 'capital controversy' (Harcourt, 1972), that is whether an aggregate measure of capital can be derived, and whether a negative relationship between the stock of capital and the rate of interest can be derived (to correspond with that postulated by the marginal productivity theory, which is retained by EGT).

The third specifically concerns notions such as 'a stock of knowledge (ideas)'. Steedman (2003) states that 'it is immediately apparent that this is far from being a crystal-clear concept. Is knowledge a homogenous quantity of which there is simply more or less? Clearly not'. As Metcalfe has argued, 'Part of the problem here is in conceiving of an aggregate stock of ideas. Are ideas to be added, multiplied together, or aggregated in combinatorial fashion, in which case the stock grows faster than exponentially? Whatever the process of aggregation, we still need the weights (prices) with which an idea in carbon chemistry, say, is to be combined with an idea in the production of insurance services. It is not obvious what these weights are, and they certainly are not to be found in market prices' (Metcalfe, 2001, p. 580).[3]

Steedman continues by noting that even if 'knowledge' can be rendered homogenous, 'the question arises whether there exists any *cardinal* measure of the single stock of knowledge' (p. 127). Labelling the stock of knowledge A, as Steedman argues, it is often assets that 'a function with A as one of its arguments does (or does not) exhibit constant returns to scale' or that 'in such a function, A has a decreasing (or increasing) marginal product' (p. 128).

NATURE OF INVESTMENT AND PRODUCTIVITY GROWTH

The particular issues raised by investment are:

(i) investment is equal to savings (through adjustment, presumably, of the rate of interest) and the savings rate is determined by individual preferences. There is no role for corporate savings, other than as proxy for individuals;
(ii) private investment comes from the marginal product of capital being equal to the rate of interest, and the marginal product of capital is based on future prices. There is no role for capacity utilisation, profitability and expectations of future demand;
(iii) hence, there is no role for aggregate demand and a form of Say's Law is assumed to operate.

Changes in the capital stock (however conceptualised) arises from economic decision-making, and the question arises as to the factors which influence those decisions. Clearly some form of production function may exist (in the sense that output is related with inputs), but what determines the inputs? The EGT answer is presumably relative prices and profit calculations. The Keynesian answer would be demand. Also the rate of technical change, however calculated, comes to depend on demand.

In EGT, the rate of capital accumulation is governed by the rate of savings, and all savings are invested. Hence the rate of investment does not ultimately depend on decisions made by firms since that rate has to conform to the rate of savings. Consider the model where firm's output depends on its capital stock, aggregate capital stock and its labour force. Because of the externality whereby its capital stock impacts on the output of other firms, the individual firm equates the marginal product of its own capital stock to the rate of interest, and its private investment decisions are socially non-optimal, which would require equating marginal social product of capital with rate of interest.

At one level, the externality effect points in the direction of government intervention along 'market failure' lines to correct the externality. At another level, the externality effect does not matter because the level of capital accumulation will be unaffected, 'since this private marginal product falls short of the average product, growth is too low in decentralized equilibrium' (Barro and Sala-i-Martin, 1995, p. 149). The key variable which influences investment decisions is the rate of interest, with the marginal product of capital equated to the rate of interest. The future is deemed to be well known to the firms, but then how this sits with an analysis of technological change remains a mystery. It would seem to be implicitly assumed that there is no problem relating to future demand. The relevant markets will clear in the future at prices which are already known to the firms, and those prices enter into the (discounted) profit calculations (see, also, chapter 6 of Barro and Sala-i-Martin, 1995, for the imperfect competition case).

As for other views on technological change, in the Barro and Sala-i-Martin (1995) approach technological change comes in two ways. The first (pp. 146–52) relates to 'learning by doing' (as they initially refer to it and as it is indexed) but it is really 'learning by investing'. This is then represented by a production function at the level of the firm i of the form:

$$Y_i = F(K_i, A_i L_i) = F(K_i, KL_i)$$

where K is the aggregate capital stock and constant returns. On aggregation, this provides a function that has constant returns to capital and increasing returns to capital and labour taken together.

The second (chapter 6) concerns research and development with respect to new products: 'a realistic description of this research process would include

uncertainty about the quantity of resources required to generate an invention and about the success of the invention. We simplify the analysis, however, by assuming that it takes a determinate amount of effort to generate a successful new product' (p. 2), though they indicate that the issue of uncertainty is taken up in the chapter that follows. A monopolist seeks to develop a new product, and this generates a production function of the form (a Cobb–Douglas production function is used here):

$$Y = A^{1/(1-\alpha)}\alpha^{2\alpha/(1-\alpha)}.LN$$

where N is the number of products, which could be said to have an AK-type form. Note that 'we assume, as usual, that households maximise utility over an infinite horizon' (p. 218). It could be said that in the endogenous growth theory, all productivity increases are 'paid' for and that there is no 'manna from heaven'. It is the growth of capital which enables labour productivity to increase. No 'manna from heaven' may sound plausible, but rules out 'learning by doing' and 'injections' of new ideas from outside the profit sector.

The Kaldorian approach involves technical progress coming through a variety of routes but here we mention two. The first is a vintage of capital effect and the second concerns economies of scale which include forms of learning by doing effects. Verdoorn's law, whereby the rate of productivity growth is a function of the rate of output growth, is a summary version of this. A feature of 'learning by doing' is that productivity gains come from the process of production, and there is an element of 'free good' involved. The implementation of the 'learning by doing' may involve resources, and different institutional structures are more or less conducive to 'learning by doing'. The point to be made here is that 'learning by doing' effects are not readily measured and are overlooked in the standard regressions involving growth and human capital.

EMPIRICAL EVIDENCE

Although the EGT is often derived in a mathematically sophisticated manner, there has been little (if any) testing of any precise propositions, which can be derived from the theory (or its variants). There has, of course, been something of an industry devoted to the estimation of growth equations, notably cross-country growth equations in which (per capita) growth has been related to a wide range of variables. The major ingredient of the variables which have been included is that they can in some way be said to relate to (growth of) capital, whether human capital (as reflected in education, health status), infrastructure, social capital (including measures of corruption, democracy) and intellectual capital (research and development etc.).

The empirical contributions to EGT are rather controversial and inconclusive. The most popular, and influential, empirical framework employed to test EGT has been cross-country regressions à la Barro (1991) and Barro and Sala-i-Martin (1995). Various studies that have used this empirical framework suggest that factors such as educational attainment, life expectancy, investment to GDP ratio, government spending to GDP ratio, financial sector development, and political stability are important determinants of long-run growth (Barro and Sala-i-Martin, 1995; King and Levine, 1993; Benhabib and Spiegel, 1994, are early examples). It would appear that such results provide comprehensive empirical backing to EGT. However, it is difficult to draw any conclusions on EGT from cross-country growth regressions for a number of reasons. The first is that it is never clear whether the variables utilised in the empirical framework are proxying for differences in country steady-state income levels or for differences in country long-run growth rates. Secondly, evidence on convergence shows that countries converge to their steady-state; but this has little bearing on the empirical verification of the EGT since it guides us poorly on predictions concerning income convergence *across* countries.

Another empirical approach, also based on cross-section growth regressions, relies heavily on the nature of returns to capital using either a Cobb–Douglas production function or an augmented Cobb–Douglas production function with human capital. The evidence gauged from these studies obtains estimates of output elasticity of less than one with respect to capital. However, there is the argument that imposing the assumption of homogeneous rate of technological progress across countries in cross-section studies is likely to lead to substantial bias in the estimates of the returns to physical and human capital (see Lee, Pesaran and Smith, 1997, 1998). Further problems with cross-section growth regressions range from econometric problems (Temple, 1999), problems of economic interpretation (Pack, 1994), to causality issues (Aghion and Howitt, 1998; Arestis an Demetriades, 1997).

Not only are there serious problems with studies that utilise cross-section country regressions, but also with those studies that have attempted to employ time-series data. The most serious problem with the time-series studies is that they are heavily inconclusive. Bean (1990) and Kocherlakota and Yi (1997) find evidence that is consistent with the EGT. By contrast, Jones (1995) and Evans (1997) provide evidence against it. Lau and Sin (1997), when focusing on the relationship between output and physical capital, are only able to provide inconclusive evidence.

It could be said that the discovery of robust relationships between economic growth and some key variables has not happened. These empirical works run into severe problems of reverse causality, since economic growth may itself be important influences on education, health etc. The empirical work has then focused on the relationship between output (growth of) and a range of inputs

(growth of). When we consider variables such as education, it would be rather surprising if no relationship was found between growth and those variables, and hence the finding of some positive relationship comes as no surprise. But the question here is: what does such a relationship tell us? Leaving aside the question of causation for the moment, the key question to our mind is what are the determinants of the growth of the relevant variables. Little has been forthcoming from the EGT proponents on this issue. The significant property of capital (as a general term) is that it involves the creation of 'assets' in the present through the use of current resources, which generate (it is hoped) future benefits. Decisions have to be made on the creation of capital, and there are obviously different views on what the key determinants might be.

In this context, there are two that stand out. The first is that of the role of actual and expected future aggregate demand. There is clearly much evidence to the effect that investment in fixed capital formation is strongly influenced by variables that are related to the level of demand (such as capacity utilisation, profitability, liquidity). In the area of research and development there has been a long debate over the roles of demand-pull and technology-push. The second is the role of government. In the EGT the role of government is limited to the 'market failure' role of providing 'public goods' (admittedly widely defined). Any notion of the government promoting industrialisation, or being a developmental state, does not arise (except perhaps in a negative sense of protectionism).

Regressing (say) output on inputs tells us nothing about the causes of the change in inputs, and the effect which demand may have had in 'bringing forth' those inputs, e.g. growth of capital stock, or the movement of people into or out of the work force. León-Ledesma and Thirlwall (2002) suggest that 'the implication for growth theory is that it makes little economic sense to think of growth as supply constrained if, within limits, demand can create its own supply. If factor inputs (including productivity growth) react endogenously, the process of growth, and growth rates differences between countries, can only be properly understood in terms of differences in the strength of demand, and constraints on demand. That is not to say, of course, that input growth is not important for output growth, but it is not *causal* in the neoclassical sense. Demand constraints are also likely to be related to supply bottlenecks which cause inflation and balance of payments difficulties for countries. It is this aspect of supply, and not the growth of inputs in a production function, that should be the main focus of enquiry in any supply-orientated theory of economic growth' (p. 456).

Furthermore, León-Ledesma and Thirlwall (2002) argue that in the relationship

$$DU = a_1 - b_1(g)$$

when $DU = 0$, this gives as estimate of the natural rate of growth the ratio a_1/b_1. Also estimates for:

$$g = c + dDU$$

include a dummy for years when actual growth is above natural rate. The dummy is estimated uniformly positive, so that 'if we go back to the Cobb–Douglas production function […] it is easy to see how it can be used for analysing the sources of growth; that is, decomposing a country's growth rate into the contribution of capital, labour and technical progress. The question is, how useful is it for a proper *understanding* of the growth performance of countries if the main inputs into the growth process are not exogenous but *endogenous*?' (Thirlwall, 2002, p. 25). We may, thus, conclude that it is rather difficult to argue that the evidence favours the EGT.

SUMMARY AND CONCLUSIONS

Much, though by no means all, of the literature on growth has referred to a closed economy, or at least with little if any reference to trade and foreign investment considerations. As far as we are aware, the neo-classical models developed from Solow (1994) made no reference to open economy considerations. The endogenous growth literature has to some degree concerned itself with this aspect, but to a very limited extent. The role of foreign direct investment and the role of foreign competition and import of 'new ideas' have been the main open economy issues given some consideration. But Thirlwall's (1997 is a good example) approach on the balance of payments being a serious constraint on growth has been a consistent critique of this omission and has produced important contributions to the debate.

In what sense, then, is the EGT new? A number of contributors have reasonably argued that most of the ideas deployed are not new and are to be found in the old neoclassical growth theory or other schools of thought. The originality of these studies lies more in bringing these ideas to the fore and packaging them in the most advanced form of mathematical models. At a more mundane level, the claim to newness resides, at least terminologically, in substituting 'endogenous' for 'exogenous'. What was previously taken as given is now explained. It is crucial to recognise, however, that this is a simple moving of the explanatory boundaries outwards although, as will be seen, there is also some countervailing shrinkage in deference to model tractability (Fine 2000, p. 248).

We may conclude this chapter by quoting at length Thirlwall to support the point: 'We have argued that there is nothing particularly new about "new"

growth theory or endogenous growth theory. Economists have been stressing for years the externalities associated with investment in human capital and research and development, and the possibility of increasing returns leading to tendencies towards polarization in the world economy. Finding divergence of per capita incomes in the world economy is consistent with "new" growth theory, but finding conditional convergence (holding other factors constant) is not necessarily support for the conventional neoclassical growth model, because convergence might have to do with catch-up and not diminishing returns to capital: the two issues are conceptually distinct. In the empirical studies attempting to explain growth rate differences between countries only four variables are robust: the savings-investment ratio; population growth; secondary school enrolment rate and the initial level of PCY [per capita income]. We find it surprising that so few studies examine the role of trade. We find that when export growth is included as an explanatory variable, it is a significant determinant of economic performance. We also find the debt–service ratio important in developing countries, and that inflation (at least, mild inflation) is not the enemy of growth as is sometimes claimed' (Thirlwall and Sanna, 1996, p. 151).

NOTES

1. The growth of employment is generally assumed to be in line with the growth of the labour force, and hence a constant rate of (un)employment. As noted in the quote from Dutt in the text, in the simple AK model there is no explicit reference to full employment. In other models, full employment would be deemed to result from flexibility of real wages bringing demand for and supply of labour into balance.
2. Quoted in Dutt (2003), p. 76.
3. The second half of this quote is used by Steedman, immediately after the quote from him given in the text.

REFERENCES

Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, Cambridge, MA: MIT Press.

Arestis, P. and Demetriades, O.P. (1997), 'Financial development and economic growth: assessing the evidence', *Economic Journal*, **107**(442), 783–99.

Barro, R.J. (1991), 'Economic growth in a cross section of countries', *Quarterly Journal of Economics*, **106**(3), 407–33.

Barro, R. and Sala-i-Martin (1995), *Economic Growth*, New York: McGraw Hill.

Basu, S. and Fernald, J.G. (1997), 'Returns to scale in U.S. production: estimates and implications', *Journal of Political Economy*, **105**(2), 249–83.

Bean, C.R. (1990), 'Endogenous growth and the procyclical behaviour of productivity', *European Economic Review*, **34**(3), 355–66.

Benhabib, J. and Spiegel, M.M. (1994), 'The role of human capital in economic develop-

ment: evidence from aggregate cross-country data', *Journal of Monetary Economics*, **34**(1), 143–73.
Dutt, A.K. (2003), 'New growth theory, effective demand, and post-Keynesian dynamics', in Salvadori (2003), pp. 67–100.
Evans, P. (1997), 'Government consumption and growth', *Economic Inquiry*, **35**(2), 209–17.
Fine, B. (2000), 'Endogenous growth theory: a critical assessment', *Cambridge Journal of Economics*, **24**(2), 245–65.
Harcourt, G.C. (1972), *Some Cambridge Controversies in the Theory of Capital*, Cambridge: Cambridge University Press.
Hussein, K. and Thirlwall, A.P. (2000), 'The *AK* model of "new" growth theory is the Harrod–Domar growth equation: investment and growth revisited', *Journal of Post Keynesian Economics*, **22**(3), Spring, 427–35.
Jones, C. (1995), 'Time series test of endogenous growth models', *Quarterly Journal of Economics*, **110**(3), 495–525.
Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the United Kingdom*, Cambridge: Cambridge University Press.
King, R.G. and Levine, R. (1993), 'Finance and growth: Schumpeter might be right', *Quarterly Journal of Economics*, **108**(4), 717–37.
Kocherlakota, N.R. and Yi, K.-M. (1997), 'Is there endogenous long-run growth? Evidence from the United States and the United Kingdom', *Journal of Money, Credit and Banking*, **29**(2), 235–60.
Krugman, P. (1989), 'Differences in income elasticities and trends in real exchange rates', *European Economic Review*, **33**, 1001–46.
Lau, S.-H.P. and Sin, C.-Y. (1997), 'Observational equivalence and a stochastic cointegration test of the neoclassical and Romer's increasing returns models', *Economic Modelling*, **14**(1), 39–60.
Lee, K., Pesaran, M.H. and Smith, R. (1997), 'Growth and convergence in a multi-country stochastic Solow model', *Journal of Applied Econometrics*, **12**(4), 357–92.
Lee, K., Pesaran, M.H. and Smith, R. (1998), 'Growth empirics: a panel data approach', *Quarterly Journal of Economics*, **113**(3), 319–23.
León-Ledesma, M.A. and Thirlwall, A.P. (2002), 'The endogeneity of the natural rate of growth', *Cambridge Journal of Economics*, **26**(1), 441–59.
Metcalfe, J.S. (2001), 'Institutions and progress', *Industrial and Corporate Change* **10**(3), 561–85.
Oulton, N. (1996), 'Increasing returns and externalities in UK manufacturing: myth or reality?', *Journal of Industrial Economics*, **44**(1), 99–113.
Pack, H. (1994), 'Endogenous growth theory: intellectual appeal and empirical shortcomings', *Journal of Economic Literature*, **8**(1), 55–72.
Panico, C. (2003), 'Old and new growth theories: what role for aggregate demand?', in Salvadori (2003).
Romer, P.M. (1986), 'Increasing returns and long-run growth', *Journal of Political Economy*, **94**(4), 1002–37.
Salvadori, N. (ed.) (2003), *Old and New Growth Theories: An Assessment*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Setterfield, M. (1994), 'Recent developments in growth theory: a post Keynesian view', in P. Davidson and J. Kregel (eds), *Employment, Growth and Finance*, Aldershot, UK and Brookfield, USA: Edward Elgar, pp. 112–27.
Solow, R.(1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70**(1), 65–94.

Solow, R. (1994), 'Perspectives on growth theory', *Journal of Economic Perspectives*, **8**(1), 45–54.
Solow, R. (1997), *Learning from 'Learning by Doing': Lessons for Economic Growth*, Stanford, CA: Stanford University Press.
Steedman, I. (2003), 'On "measuring" knowledge in new (endogenous) growth theory', in Salvadori (2003), pp. 127–33.
Swann, T. (1956), 'Economic growth and capital accumulation', *Economic Record*, **32** (November), 334–61.
Temple, J. (1999), 'The new growth evidence', *Journal of Economic Literature*, **37**(1), 112–56.
Thirlwall, A.P. (1991/2), 'Professor Krugman's 45-degree rule', *Journal of Post Keynesian Economics*, **14**(1), 23–8.
Thirlwall, A.P. (1997), 'Reflections on the concept of balance-of-payments-constrained growth', *Journal of Post Keynesian Economics*, **19**(3), 377–86.
Thirlwall, A.P. (2002), *The Nature of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Thirlwall, A.P. (2003), '"Old" thoughts on "new" growth theory', in Salvadori (2003), pp. 44–52.
Thirlwall, A.P. and Sanna, G. (1996), '"New" growth theory and the macro-determinants of growth: an evaluation and further evidence', in P. Arestis (ed.), *Employment, Economic Growth and the Tyranny of the Market*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 131–56.

8. The nature of economic growth and the neoclassical approach: more questions than answers?

John McCombie

INTRODUCTION

The study of economic growth is back in fashion. Mankiw (1995, p. 275) has commented 'after many years of neglect, these questions [about economic growth] are again at the centre of macroeconomic research and teaching. […] Moreover, growth is not just important. It is also a topic which macroeconomists, with their crude aggregate models, have something useful to say'. Of course, concern with the determinants of economic growth, and why some countries are rich and others poor, goes back a long time; at least to Adam Smith's *Wealth of Nations* (1776), if not before. Concern with problems of development has never has really been off the economists' agenda. But Mankiw is referring here solely to neoclassical growth theory and there has indeed been a revival since the mid-1980s, with renewed emphasis on the augmented Solow growth model and also the development of endogenous growth models. Parallel with this has been the rapid increase in Barro-type regression analyses drawing on the greater availability of data, such as the *Penn World Tables*. Whether or not neoclassical growth theory has led to a greater understanding of the nature of economic growth is the question that I shall address in this chapter.

While there have been disagreements within the neoclassical framework (such as the convergence debate), Mankiw's comments accurately represent the consensus amongst neoclassical economists. Expositions of neoclassical growth theory are now central to many macroeconomic textbooks as a theory of the long run (e.g., Romer, 2001; Mankiw, 2002) rather than being tacked on at the end, almost as an afterthought, as they were in the past (e.g., Levacic and Rebman, 1982).

However, this gives a misleading impression, as there are many who are dissatisfied with the neoclassical approach. Tony Thirlwall has long been a forceful critic and a recent succinct statement of his views is to be found in his 2002 book, *The Nature of Economic Growth.* There, he notes that these neoclassical

models are 'supply-orientated, supply-driven, closed economy models unsuitable for the analysis of open economies in which foreign exchange is invariably a scarce resource acting to constrain the growth process' (Thirlwall, 2002, p. 28).

He also expresses surprise that the 'new' endogenous growth theory should be seen as new, because the implications of this approach had been emphasised many years ago by Prebisch, Hirschman, Myrdal, Arrow, and Kaldor, amongst others (see Toner, 1999). Moreover, Frankel, as long ago as 1962, developed what may be considered as one of the first formal endogenous growth models. Arestis and Sawyer in this volume raise some criticisms about the endogenous 'linear-in-capital' or *AK* model and Fine (2000) finds the insights of endogenous growth theory limited, *inter alia*, by its excessive formalism.

Nelson (1998) divides growth theory into *appreciative* and *formal theorising*. The former tends to be empirical and inductive, while the latter is deductive and builds theoretical models largely based on a few stylized facts. While formal theorizing should be a useful check on the internal consistency of the arguments in appreciative theorizing, Nelson considers that it has not really led to any insights that were not known as long ago as the 1950s (e.g., Abramovitz 1952). Kenny and Williams (2001), in a wide-ranging assessment of what has been learned from neoclassical growth theory, are not impressed. They are critical of the substantial empirical literature that has arisen in the last two decades, partly because of the lack of robustness in Barro-type regressions (Levine and Renelt, 1992), and partly because the models implicitly assume ergodicity. They contend that growth is path-dependent, with small differences in initial conditions leading to vicious and virtuous circles of growth with very different outcomes. Like Nelson, they conclude that neoclassical growth theory has not told us much that we didn't already know; except that they consider that it is all to be found in Arthur Lewis's (1955) *Theory of Economic Growth*. Ruttan (1997) also finds the approach of the old and new growth theories to modelling technical change superficial, especially when compared with path-dependency models, evolutionary growth theory, and microeconomic models of induced technical change.

In this chapter, I shall present a further assessment of the neoclassical approach to growth. Given the vast literature that has developed and, for reasons of space, this essay will necessarily have to be eclectic and will not, for example, comprehensively review the vast number of econometric studies that now exist. Furthermore, I shall not be concerned with Tony Thirlwall's influential balance-of-payments constrained growth model, which, perhaps not surprisingly, in my opinion presents a convincing alternative demand-oriented approach to the neoclassical model (Thirlwall, 1979; McCombie and Thirlwall, 1994 and 2004). Nor shall I consider the cumulative causation models of economic growth and the Verdoorn law. (See McCombie, Pugno and Soro, 2002.)

IS CAPITAL ACCUMULATION THE SECRET OF ECONOMIC GROWTH?

So where does one begin in trying to explain the vast differences in living standards across the world? One of the most apparent indicators of the level of economic development is the capital–labour ratio. The advanced countries have high per capita levels of both social overhead capital and private physical capital (i.e., equipment and structures) compared with the less developed countries. This led to the early conjecture that the key to economic growth was the rate of physical capital accumulation (Rostow, 1960).

The beginnings of modern growth theory may be taken to be the Harrod–Domar model, which Harrod (1939) saw as an extension of the Keynesian model to the long run, although Domar (1946) saw it more as a model of short-run fluctuations. In this approach, the rate of capital accumulation is paramount, as, assuming a fixed-coefficients production function, economic growth is given by the simple relationship $\hat{Y}_A = s/\nu = \lambda + \hat{L}$ where $\hat{Y}_A$ is actual output growth; s is the savings-output ratio; ν is the (constant) incremental capital–output ratio; λ is the rate of technical progress; and $\hat{L}$ is the growth of employment.[1] The warranted growth rate $\hat{Y}_W$ is the growth rate where planned savings equals investment and is the equilibrium growth rate. Central to Harrod's analysis was the instability, or knife-edge, problem, which occurred when the actual and warranted growth rates diverged. Nevertheless, we shall ignore this problem and assume that the warranted and actual growth rates are equal. The natural (or maximum) rate of growth is given by $\hat{Y}_N = \lambda + \hat{N}$, where $\hat{N}$ is the growth of the labour force. The long-run problem is to ensure that the warranted and natural growth rates are equal and there is nothing in the model that ensures this. This problem was solved by Solow (1956) and Swan (1956), who, using a production function where there was the possibility of substitution between the factors of production, showed how changes in the capital–output ratio would ensure that the warranted and actual equalled the natural or potential growth rate. The other, or knife-edge, problem was 'solved' by dispensing with a separate investment function. Consequently, now many macroeconomics textbooks (for example, Mankiw, 2002, and Romer, 2001) take the beginning of modern growth theory to be the Solow model.

Nevertheless Easterly (1997, 2001) documents how the Harrod–Domar model has been used by development agencies (including the World Bank) for much of the post-war period, and long after it had been abandoned in academic circles. It was used to calculate the 'savings gap' and the associated 'financing gap', which arose when the actual growth rate was below the natural growth rate. If the incremental capital–output ratio is known, then the ratio of investment to output needed to achieve a certain rate of growth can be calculated. Knowledge of the required investment ratio, given the domestic savings ratio, was seen as

necessary to determine the amount of foreign aid required to achieve a target growth rate for any particular developing country. Of course, it is not necessary to adopt the Harrod–Domar model to believe that capital accumulation is the key to economic growth. Easterly and Levine (2001, p. 178, fn. 2) document the widespread current belief of development economists and agencies in the central role of capital accumulation in promoting economic development and the modern *AK* growth models encourage such a belief.

The importance of the investment–output ratio in determining economic growth received early support from Hill (1964), who ran some simple cross-country regressions between the growth of GNP and the investment–output ratio for the advanced countries for the period between 1953 and 1962. He found that there was a significant relationship between growth and the share of investment in machinery in GDP, but not investment in structures (construction). Nearly thirty years later, DeLong and Summers (1991, 1992) rediscovered the importance of the investment–output ratio in growth. They found very similar results to Hill using a larger sample of countries and more recent data, and interpreted their results as implying there is a substantial externality to investment. (Auerbach *et al*., (1994) questioned the robustness of these results, but see DeLong and Summers's (1994) rejoinder). Hussein and Thirlwall (2000) present further empirical evidence, but their results are rather more mixed. Blomstrom *et al*., (1996) using causality tests find that a faster rate of GDP growth 'causes' a higher investment-output ratio and not vice versa. The implications are that investment is not a key determining variable in the growth process. Once growth is underway, this will cause the investment rate to increase, in a Kaldorian manner.

These results, however, stand in marked contrast to those obtained by Easterly (1997, 2001) and Easterly and Levine (2001) who, using time-series data for 146 countries, found no significant correlation between the growth of GDP and the investment–output ratio. Easterly (1997, p. 24) concludes that 'the Harrod–Domar growth model makes no sense theoretically and it fails empirically'. The key to growth, according to Easterly (2001) is technology, and this is discussed below.

One interpretation is that the gross investment–output ratio is merely acting as a proxy for the growth of the capital stock (but see Scott, 1999 for an alternative interpretation). However, Pritchett (1996a) has pointed out that capital–output ratios vary considerably between countries and so countries with the same investment–output ratios can have very different growth rates of the capital stock. For example, he cites the case of Korea which, with an investment–output ratio of 21.5 per cent, had a growth rate of the capital stock of 12.5 per cent per annum, whereas Germany, with a slightly higher investment–output ratio, had a growth of capital that was only one-third as large, namely 3.9 per cent per annum. Thus, the investment–output ratio is a poor

proxy for the growth of the capital stock. This suggests that the significant results from estimating the Harrod–Domar model using the gross investment ratio may well be spurious.

More worryingly, Pritchett (2000) argues convincingly that there are problems with even using data for the growth of the capital stock for many less developed countries. He contends that there is no point in cumulating depreciated investment by the perpetual inventory method (what he terms CUDIE – cumulated, depreciated, investment effort) to give estimates of, especially, the public capital stock. Piecing together indirect evidence from a number of sources, Pritchett concludes that the inefficiency of public investment could be substantial for a large number of developing countries. The implications are serious. It means that just because public spending has a small effect on growth, this does not mean that public capital is not *potentially* productive. Moreover, as a result, estimates of the growth of total factor productivity for, especially, the developing countries are so suspect as to be meaningless. Indeed, this may well account for the fact that many less developed countries record zero or negative total factor productivity growth, as the growth of capital estimated by CUDIE considerably overstates the true growth of the effective capital input. This raises serious problems for both policy and for cross-country Barro-type and other regressions that use the growth of capital. Furthermore, Kenny and Williams (2001) note that a number of studies show weak correlations between different estimates of the same variables (for example, GDP or exports) produced by different agencies or found in different sources (World Bank, *Penn World Tables*, IMF, and so on). These measurement problems are likely to be so serious that instrumental variable approaches are unlikely to solve the difficulties that they pose for econometric estimation. Disraeli's aphorism about 'lies, damn lies and statistics' certainly seems applicable here.

THE SOLOW GROWTH MODEL – IS TECHNICAL CHANGE AND HUMAN CAPITAL THE ANSWER?

The Solow growth model has long been the workhorse of macroeconomic growth modelling. While it has been extended by the endogenous growth models, it is nevertheless still the benchmark against which other models are judged. As is well known, the Solow growth model came up with two surprising predictions. First, the share of investment in output does not affect the steady-state rate of growth, which is determined solely by the rate of technical change. It merely affects the steady-state *level* of productivity. Thus, contrary to the Harrod–Domar model, increasing the share of investment does not lead to a sustained increased growth of productivity, because of the presence of diminishing returns to capital. Secondly, by far the largest contribution to the growth of

output and productivity was, with one or two notable exceptions (Young, 1992, 1995), found to come from technical change, rather than the growth of the factor inputs. As there was initially no satisfactory theory of the determinants of technical change, this was simply treated as exogenous (Solow, 1957). The key concept of this analysis is the aggregate production function, about which I shall have more to say below.

At first glance, the Solow model, with its reliance on the aggregate production function and the assumption that all countries have access to the same technology, does not perform well. Romer (1994) takes a simple example, comparing the Philippines with the US. Given their respective capital–labour ratios and abstracting from differences in human capital, the US would have to save about 30 times per capita as much as the Philippines for the two countries to grow at the same rate. (Their growth rates do not greatly differ.) In fact, there is not much difference in their savings rates. Prescott (1998), doing a related exercise, is equally sceptical. Lucas (1990), using a 'back-of-an-envelope' calculation, suggests that under standard Solow assumptions, the marginal product of capital should be 58 times higher in India than the US because of the former's lower capital–labour ratio. In fact, the difference is nowhere near as great as this. Nearly all portfolio investment and skilled labour flows to the advanced countries, not to those with low capital–labour ratios, whereas if Solow was right, one would expect the flows to be the other way. Given the shortage of skilled labour in the less developed countries, its marginal product should be much higher than in the advanced countries.

In the light of all this, it is tempting to conclude that the Solow model simply cannot explain the observed wide differences in per capita income. Furthermore, the model also predicts convergence of per capita income and the evidence suggests that for the world as a whole there is no significant relationship between the growth of output per capita and the initial level of output per head. This stylized fact, if it can be called that, was one of the reasons for the development of the endogenous growth models that can predict divergence.[2] This led Prescott (1998) to suggest that what is needed is a theory of total factor productivity, which cannot be taken to be constant across countries.

Nevertheless, Mankiw (1995) still finds the Solow model convincing, and the paper by Mankiw, Romer and Weil (1992, p. 407) 'takes Robert Solow seriously'. Mankiw (1995) explains away Lucas's argument by pointing out that the differences in the rate of return between countries are greatly diminished if it is assumed that the elasticity of substitution between labour and capital is high, say, four or greater. He squares this with the stylized fact that factor shares are constant, which would normally suggest that the production function is a Cobb–Douglas, by arguing that this occurs because countries (or industries) are growing at their steady-state growth rates. This, regardless of the exact value of the aggregate elasticity of substitution, will give constant factor shares.

But this makes the hypothesis untestable.[3] If Mankiw is correct and all industries are at their steady-state growth rate, then the elasticity of substitution cannot be estimated. It could be equally zero (fixed coefficients) as four or, indeed, ten. In fact, the estimates from production function studies suggest that if anything the elasticity of substitution is *less* than unity (see Rowthorn, 1999).[4] Mankiw further suggests that the elasticity of substitution could be very high if, as capital accumulation occurs, there is an increased specialisation in capital-intensive exports. 'In other words, as international trade works to equalize factor prices around the world, it drives the effective elasticity of substitution in each economy to infinity' (Mankiw, 1995, p. 288). However, the factor price equalization theory is probably one of the most decisively refuted hypotheses in economics.

Mankiw (1995, p. 295) accounts for the 'perverse' migration of skilled labour by noting that 'because human and physical capital are *complementary* inputs in production, imperfections in the financing of human capital impede the international movement of physical capital' (emphasis added). But this requires a very different production function to the Cobb–Douglas or any of the other standard production functions (CES, translog, etc.) used in the augmented Solow growth model. An interesting approach is Kremer's (1993) O-ring theory of economic development. (The O-ring was a small component of the Challenger space shuttle that failed on launch, resulting in disaster.) This sees a worker with a particular skill as affecting the productivity of other co-workers with different skills. The labour input can, for example, be modelled as $(\prod_i q_i)L$ where q_i is an index of the relative level of skill of a worker and L is the number of employees. Because of the multiplicative nature of the skills variable, a low-skilled worker can dramatically reduce the efficiency of all the other workers and hence, aggregate productivity.[5] Consequently, if most workers in a less developed country are low skilled, the fact that a skilled worker needs to co-operate with them in production means that the skilled worker's marginal product will be well below that obtained in an advanced country. This requires a very different production function to the ones that are currently employed in neoclassical growth models.

Mankiw (1995) defends the assumption that all countries share a common technology. If all the difference in productivity levels was due to disparities in the levels of technology then 'the incentives to imitate technology used by rich countries would be tremendous' (p. 283), although he does concede that this would require an appropriately skilled labour force. Others are not convinced. Nelson and Pack (1999) argue that new technology cannot just be taken down from the shelf, as it were, and used. There is a substantial amount of learning and adaptation required. There is a great deal of evidence to suggest that technology transfer is not instantaneous but is a function of, *inter alia*, the degree of social capability of the labour force (Abramovitz, 1986). (See also Nelson,

1991; Fagerberg, 1994; and Nelson and Wright, 1992.) Nelson and Pack (1999) develop a two-sector model to explain the East Asian economic miracle. The growth of overall productivity is determined by the speed of transfer of workers from the traditional low-productivity sector to the modern high-productivity sector, the latter being dependent on foreign investment and technological transfer. The rate of transfer is a function of the degree of dynamism of the entrepreneurs and the growth of skills of the labour force. These important sources of growth, Nelson and Pack (1999) maintain, are concealed in the one-sector neoclassical model.

Notwithstanding these arguments, Mankiw suggests that we should assume that all countries are on the same production function and focus on differences in education. If we introduce human capital it is found that capital's share, broadly defined, increases from one-third to two-thirds, and, according to Mankiw, most of the disparities between the advanced and the less developed countries can be explained. If we then allow for the fact that growth rates may differ because countries are not at their steady-state growth rates, we have, according Mankiw, the beginnings of a convincing explanation of why growth rates differ (Mankiw, Romer and Weil, 1992). Assuming an aggregate Cobb–Douglas production function, constant returns to scale and fully-employed resources, and that each country from, say, Guyana to the US,[6] has access to the same blueprint of technology and the same rate of technical progress (λ), Mankiw, Romer and Weil (MRW) show that the steady-state Solow model augmented by human capital can be expressed using cross-country data as:

$$\ln y = \ln A + \frac{\alpha}{1-\alpha-\beta}\ln s_K + \frac{\beta}{1-\alpha-\beta}\ln s_H - \frac{\alpha+\beta}{1-\alpha-\beta}(n+\delta+\lambda) \qquad (1)$$

where y is output per capita; A is the level of total factor productivity, which is assumed to be the same for all countries; s_K is the ratio of investment in physical capital to total output; s_H is the ratio of investment in human capital to total output; n is the growth of population; δ is the rate of depreciation; and λ is the common rate of technical progress.[7] The parameters α, β, and $(1 - \alpha - \beta)$ are the output elasticities of physical capital, human capital and unskilled labour. This equation is estimated by MRW using the *Penn World Tables* and cross-country data. As has been mentioned, equation (1) assumes that the countries are at their steady-state growth rates, but as this might not be the case, the initial level of productivity is also included in a subsequent specification as a regressor to estimate the speed of convergence. (This, however, contradicts Mankiw's argument noted above that the elasticity of substitution could greatly exceed unity even though factor shares are constant, because all countries are in steady-state growth.) By and large, they find the results convincing. The estimates of each of the output elasticities are 1/3 and the estimated speed of convergence

is around 2 per cent per year, which is compatible with these values. Generally, the statistical fit of the augmented Solow model is reasonably good, although puzzlingly, this is not true for the sub-sample of OECD countries. The implication is that the automatic convergence of countries' per capita incomes to their steady-state levels is only a matter a time.

But there are a number of severe shortcomings with this model and hence with the results of estimating the Mankiw–Romer–Weil model. Some of these problems are well known, but are no less important for that, and it is worth rehearsing them.[8]

While the model may be able to explain conditional convergence, the history of the past two centuries has been 'divergence, big time' to use Pritchett's (1997) evocative phrase. Two hundred years ago there was very little difference between the per capita income levels in different parts of the world. Some estimates suggest that the ratio was about 1:3. Today, the ratio is of the order of 1:50, if not larger. Thus, the Solow model has difficulty in explaining why disparate growth rates and per capita income levels occurred in the first place, apart from postulating country-specific shocks.

Is it plausible to assume, as the model does, that perfect competition prevails in all countries and population growth, especially in less developed countries, is an adequate proxy for the flow of labour services? There is considerable evidence of disguised unemployment in many developing (and advanced) countries. Moreover, the efficiency of the labour force in the less developed countries is adversely affected by malnutrition, lack of medical facilities, and a greater propensity for illness. Thus, the increased flow of labour services resulting from a given growth in the population will be considerably less in the less developed countries than in the advanced nations. Indeed, it seems almost certain that for the developing countries there is no well-defined causal relationship between the growth of population and output, except possibly from the latter to the former. The growth of the effective labour input must, at the very least, be seen as endogenous, and is difficult to measure.

The savings ratio may also be a function of the capital–labour ratio with low savings rates associated with low capital–labour ratios and this can cause multiple equilibrium with a low-level equilibrium trap (Jones, 1975, p. 92; see also Nelson, 1956). Sachs *et al.* (2004) also use the neoclassical growth model when discussing Africa's poverty trap. But their arguments do not necessarily depend upon this model. They attribute Africa's lack of development to three major factors: (i) the lack of a good transport network especially for the many landlocked countries; (ii) poor agricultural productivity because of climatic reasons and the meagre use of fertilizers, which are costly to import and transport; and (iii) the very serious incidence of disease arising from malaria and, more recently, the prevalence of HIV/AIDS. They argue that Africa needs carefully targeted aid to overcome these and other obstacles to development before the

continent can escape from the poverty trap. Indeed, I suspect that there would be little disagreement about either their diagnosis or their proposed prescriptions, but these are *not* dependent on either the Solow growth model or the *AK* endogenous growth model. (The *AK* model is discussed in greater detail below.) Indeed, it is arguable that these models could have been dispensed with in the discussion without affecting the force of Sachs *et al.*'s arguments.

As we have seen, the missing explanatory variable in accounting for disparities in development, according to Mankiw, is human capital. But there are also problems with this argument. Generally, the empirical evidence does not support this conclusion when human capital is properly measured (Pritchett, 1996b). Although there was a rapid growth of school provision in many less developed countries between 1960 and 1990 and especially the African countries invested heavily in education, many were nevertheless 'growth disasters'. Thus, Easterly, (2001, p. 74) comments that 'Zambia had a slightly higher rate of expansion in human capital than Korea, but Zambia's growth rate was seven percentage points lower. Simply considering enrolment rates is very misleading; what matters is the *quality* of the education and the appropriateness of the syllabus for the level of development of the country concerned'. Both these are seriously deficient in less developed countries and also differ between the advanced countries (Todaro and Smith, 2005). Finally, there is again the question of the direction of causation. As Solow (2000) has remarked with respect to Barro regressions, it is difficult to know on which side of the equation the many variables used in the analyses should go.

These models also pay no attention to the structure of production and the type of industries in a country, both of which are crucial in explaining the level of development (Chenery and Syrquin, 1975). An increase in investment is assumed to generate output that will automatically be sold. There is no consideration of the problems facing developing countries in producing high-tech products for which productivity and world demand are growing rapidly, rather than being trapped in the production of primary products and basic manufactures for which the growth of world demand is low. Indeed, it is in fostering these industries that government has an important role to play in promoting development (Wade, 1990), which cannot be left to the market.

Finally, the models exclude international financial flows and associated currency crises that have been seen as causing, in some cases, a lost decade of growth. There is also no reference to the fact that the growth rate itself can be limited by the balance-of-payments constraint if the growth of foreign exchange revenues is not sufficiently fast (Thirlwall, 1979). The growth of exports, consequently, plays an important role in the growth process. These deficiencies with the neoclassical growth model stem from the starting assumption that economies are closed and demand responds passively to changes in supply.

THE GROWTH ACCOUNTING APPROACH

Given the problems of econometrically estimating aggregate production functions, especially using time-series data (multicollinearity, endogeneity, etc.), an alternative approach has been simply to assume that factors are paid their marginal products and, consequently, that the output elasticities are equal to the factor shares (Denison, 1967; Barro, 1999). In the 'growth accounting approach', the growth of total factor productivity (*TFPG*) is calculated using the aggregate production function in growth rate form as:

$$TFPG \equiv \hat{Y} - a_t\hat{K} - (1 - a_t)\hat{L} \tag{2}$$

where $\hat{Y}$, $\hat{K}$, and $\hat{L}$ are the growth rates of output, capital, and labour respectively, and a and $(1 - a)$ are the shares in output of capital and labour. *TFPG* is thus the growth of output less the growth of the factor inputs each weighted by its factor share, which, under the usual neoclassical assumptions, as we have noted, are taken to measure the output elasticities. The exact functional form of the aggregate production function is not required, although the assumption of Hicks-neutral technical change is needed. While this approach is subject to the criticisms noted above concerning the neoclassical approach in general, it is useful to use this framework to discuss some further conceptual problems.

A great deal of effort over the years has gone into refining this approach in terms of improving the quality of the data and introducing other factors such as the impact of government regulation, increasing returns, etc. However, the basic procedure has remained the same for the last fifty years or so; see OECD (2004) for a recent study using this approach. It largely stems from Solow's (1957) seminal paper – although the pioneering contributions of others, such as Abramovitz (1956), should not be overlooked. The conclusions found by Solow (1957) have been broadly confirmed. In his original study Solow found what he considers to be the 'startling' result that total factor productivity growth accounted for seven-eights of the growth of labour productivity growth of the non-farm private business sector of the US over the period 1909–49. Solow (1988, p. 314), commenting later on Denison's (1985) study, concluded that 'education per worker accounts for 30 percent of the increase in output per worker and the advance of knowledge accounts for 64 percent in Denison's figures. Thus, technology remains the dominant engine of growth, with human capital investment in second place'. This is 'Solow's shocker' as Easterly (2001, p. 49) puts it.

This is also confirmed by Easterly and Levine (2001, p. 138) who report the results of a number of studies and conclude total factor productivity (TFP) 'growth accounts for about half of output growth in OECD countries. Variation is greater among Latin American countries, with an average of over 30 percent'.

(See also Maddison, 1987.) Young (1992) found controversially that TFP growth in Singapore over the period 1966–90 was actually negative, but this result is an exception, and the calculations have not gone unchallenged (Hsieh, 2002). Easterly and Levine (2001) also undertake variance decomposition of output growth using the *Penn World Tables* and find 'TFP growth differentials account for the bulk of cross-country differences' (p. 187). This is further evidence that contradicts the assumption of Mankiw that all countries share the same production function. It should be noted that the differences in TFP growth in the growth accounting approach cannot be attributed to transitional dynamics. The latter occur because a country that is, for example, below its steady-state level of output per worker will have a faster growth of the capital–labour ratio, but this is explicitly taken into account in the growth accounting approach.

However, the growth accounting approach is not uncontroversial. Nelson (1973), some time ago, argued that the growth accounting approach had reached diminishing returns, if not a dead end. One of the problems is that it needs the assumption of Hicks-neutral technical change. If this does not occur, and it is difficult to think of any reasons why it should, and there is, say, biased technical change, then the growth of labour's share, which is empirically approximately zero, is given by:

$$(1\hat{-}a) = -a\left(1-\frac{1}{\sigma}\right)[(\lambda_K + \hat{K}) - (\lambda_L + \hat{L})] \approx 0 \qquad (3)$$

where a, σ, λ_K, λ_L are capital's share, the elasticity of substitution, capital-augmenting and labour-augmenting technical change. It can be seen that at any point in time the value of labour's and capital's shares are not independent of the rate of technical progress and the elasticity of substitution and so cannot be used to weight the growth of labour and capital to provide an estimate of the contribution of these factors to growth, *independent* of the rate of technical change. See Nelson and Pack (1999) and Felipe and McCombie (2001) for discussions.

There are other conceptual problems with both this approach and the estimation of neoclassical production functions, in addition to those noted above. First, as Kaldor (1957) noted a long time ago, technical change, the growth of output and the rate of capital accumulation are inextricably interrelated, and there is not much sense in trying to calculate the separate contributions of each, either through the growth accounting procedure or by estimating production functions.

Nelson (1973) uses a striking analogy of a cake to illustrate this critique. There are many ingredients that go to make up a cake, such as milk, flour, eggs, and so on. While it is possible to determine the effect of the change in the cake caused by, say, the addition of another egg, it makes no sense to ask the question,

what is the percentage contribution of the eggs to the cake? But this is exactly what the growth accounting approach attempts to do. Abramovitz (1993, p. 221) forcibly makes the same point: 'there are two-way connections between technological progress, economies of scale, tangible capital accumulation, and human and other intangible capital accumulation'. So it is not a question of just calculating the contributions to output growth of the growth of physical capital, human capital, and labour, regarding these as the proximate sources of growth and then seeking to find out what determines these – unfortunately it is more complicated than this. Nor can this be adequately modelled simply by including an interaction term in the estimation of a production function, because the interactions are too complex and change over time.

WHAT IS NEW ABOUT THE 'NEW' GROWTH THEORY?

We have seen that two shortcomings of the Solow model are, first, there is no satisfactory explanation of what determines the level and growth of total factor productivity ($A(t)$ and $\hat{A}$). Secondly, it predicts convergence, whereas the story over the last two centuries has generally been one of divergence, albeit with more recent 'convergence clubs'. It was partly to overcome these two problems that the new, or endogenous, growth models were first developed in the 1980s, although they are flexible enough to incorporate the diffusion of innovations from the advanced to the less developed countries. For succinct surveys of these models and their development, see, *inter alios*, Romer (1994) and Solow (2000).

However, the designation 'new' is misleading, because the endogenous growth models build on the early Solow and other growth models. The difference is that parameter values that were previously considered implausible are now seen by some as credible (such as the fact that there are no diminishing returns to capital) (Cesaratto, 1999).

The endogenous growth models usually take as their starting point the traditional aggregate production function, assumed, for simplicity, to be a Cobb–Douglas of the form $Y = K^{\alpha}(AL)^{(1-\alpha)}$, where Y, K, A, and L are output, capital, the level of technology, and employment. α and $(1-\alpha)$ are the output elasticities of capital and labour. The trick is to express $A(t)$ as a function of variables that are determined endogenously within the model. The models are also firmly wedded to the neoclassical tradition of concentrating on steady-state solutions. This places limits to the type of functional forms that the various models can accommodate and so crucial assumptions underlying the models are what might be termed *theory dependent.* For example, $A(t)$ must grow at a smooth exponential rate (or else be constant) and so certain functional relationships are necessary in the model for a steady-state solution, but these are

presented with little or no justification of either a theoretical or empirical nature (Solow, 2000, p. 100). Solow (2000) likens the search for the specific functional form in the various models to the game of 'Where's Waldo?' Consequently, confining our attention to the determinants of $A(t)$, we can specify a number of functional relationships that will either fully or partially endogenize growth. The difference between the fully- and semi-endogenous growth models is that the latter still need the growth of population, or strictly speaking employment, to generate steady-state growth. We may identify four functional forms for $A(t)$ that have been commonly used in the growth models.

(i) $dA/dt = \lambda A$, which implies $\hat{A} = \lambda$ and $A(t) = A_0 e^{\lambda t}$. If λ is exogenous, this is nothing more than the traditional Solow model.

(ii) $(dA/dt)/A = 0$. If there is no exogenous technical change, then for steady-state growth to occur there can be no diminishing returns to capital. If K is the capital stock, broadly defined, this provides the basis for the 'linear-in-K' model or $Y = AK$ model, where A here is a constant of proportionality.[9] The first generation models (Romer, 1986, 1987) were an extension of Arrow's learning by doing model. Learning-by-doing can be specified as a function of, for example, the growth of the capital stock. If it is assumed that each firm can gain access to the benefits of other firm's learning by doing in a costless manner then an individual firm's production function can be expressed as $Y_i = K_i^{\alpha}(AL_i)^{(1-\alpha)}$ where $A = (K/L)^{\theta}$. Thus, the economy-wide level of technology is a function of the aggregate capital-labour ratio. As the K/L ratio is an externality for the individual firm, the assumption of perfectly competitive markets may be maintained. At the economy-wide level,[10] the production function is given by $Y = K^{\psi}L^{(1-\alpha)(1-\theta)}$ where $\psi = \alpha + \theta(1 - \alpha)$. Steady-state growth with increasing productivity will only occur if $\theta = 1$ and, hence, $\psi = 1$. We now have a 'linear-in-K' production function. The steady-state growth rate is now a function of the investment-output ratio so that the growth rate is amenable to policy.

The problem with these type of models is that if $\psi < 1$, we are back in the familiar Solow world.[11] We have diminishing returns to capital, and as there is no technical change to offset it, the growth rate of productivity tends to zero with time and $\hat{Y} = \hat{K} = \hat{L}$. But this is at variance with Kaldor's (1961) stylized fact that productivity grows at a steady rate, so it is necessary to incorporate exogenous technical change again. If $\psi > 1$, there is explosive growth and it only needs ψ slightly to exceed unity to create results that are most implausible. For example, if $\psi = 1.05$, and the savings ratio is 10 per cent, 'a country like Germany or France will achieve infinite output in about 200 years, or even a shorter time from "now"' (Solow 1994, p. 50). Consequently, we have something of a knife-edge problem here. The coefficient of ψ needs to be exactly equal to unity, but

there is no reason why this should be the case, except for a remarkable coincidence. For these reasons, *inter alia*, the 'linear-in-K' model has largely been discarded.

(iii) $dA/dt = \gamma Z A^{\xi}$ which, if $\xi = 1$, implies that $\hat{A} = \gamma Z$. For this to form the foundation of an endogenous growth model, Z must be constant in steady-state growth. Lucas (1988), building upon an earlier model by Uzawa (1965), in effect, takes A as human capital (H) and $Z = (1 - u)$, the share of labour time that is spent on human capital formation instead of producing goods. The production function is assumed to take the form $Y = K^{\alpha}(uHL)^{(1-\alpha)}$. Consequently, as $dH/dt = \gamma(1 - u)H$ and, assuming H grows at a constant exponential rate, the production function can be expressed as $Y = K^{\alpha}(uH_o e^{\gamma(1-u)t}L)^{(1-\alpha)}$. This again is similar to the Solow model, but where $\lambda = \gamma(1 - u)$. The difference is that the rate of growth of human capital is assumed to be given by an optimising procedure that determines $(1 - u)$ and so λ is determined endogenously within the model. But note that the assumption that $\xi = 1$ is both a very strong one and necessary to generate steady-state growth. We have again a knife-edge assumption.

(iv) $dA/dt = \rho Z^{\pi} A^{\phi}$ which implies that in steady-state growth, $\hat{A} = (\pi/(1 - \phi))\hat{Z}$ on the assumption that $\phi < 1$. It will be noted that this functional form is very similar to that in (iii), although the values of the parameters differ. If A is interpreted as the number of 'ideas' and Z the number of workers in the R&D sector (namely, L_A), then this is the basis of Romer's (1990) model, although Romer himself assumes $\phi = 1$. This is implausible as it implies that $\hat{A} = \rho L_A^{\pi}$ where the growth rate of new ideas is a constant so long as the number of researchers does not grow. A corollary of this is that if the number of workers in the R&D sector grows over time, as empirically is the case, so the growth rate of the economies should increase over time, which has not happened (Jones, 1995).

A simplified version model of this model again starts from the standard Cobb–Douglas production function $Y = K^{\alpha}(AL_Y)^{(1-\alpha)}$ where A is the stock of 'ideas' and L_Y is the number of workers employed in producing Y. The new ideas are produced in a R&D sector. In this model the number of new ideas produced is given by $dA/dt = \rho L_A^{\pi} A^{\phi}$, where $\phi < 1$. As we have noted, this implies that the steady-state growth of the stock of new knowledge is given by $\hat{A} = \pi \hat{L}_A/(1 - \phi)$ and $\hat{L}_A = \hat{L}_Y$. The form of the production function in steady state can now be written as $Y = K^{\alpha}(AL_Y)^{(1-\alpha)}$ where $A = \exp \int \hat{A} dt$, or $Y = K^{\alpha}(A_0 e^{\lambda t}L)^{(1-\alpha)}$ where $\lambda = \pi \hat{L}_A/(1 - \phi)$ if $\hat{L}_A$ is constant. It can be seen that this has all the hallmarks of the Solow growth model, with the exception that λ can be determined by an optimisation procedure, reflecting the time preference for present versus future consumption (which reflects the ratio of L_A to L_Y). It is only partially endogenous as steady-state growth is still a function of the growth of the labour supply, which

in steady-state equals the growth of workers in the R&D sector. But as it stands, increasing the share of physical investment or the proportion of labour in the R&D sector (L_A/L) has only level and not growth rate effects. The interesting point to note is that this production function is assumed to be for the advanced countries as a whole. The less developed countries acquire 'new' ideas from the advanced countries and the rate at which this diffusion occurs depends upon their level of human capital.

It can be seen that there are striking similarities in the functional forms of (iii) and (iv). The difference is that in the Lucas model, as the share of labour time spent forming human capital is likely to be roughly constant (it is bounded by 0 and 1), the steady solution requires $\xi = 1$. However, in the Jones model, as the number of researchers grows over time, the corresponding coefficient, ϕ, must be less than unity. Hence, the values of both these parameters are theory driven and it is not hard to find Waldo.

From this short discussion, it is readily apparent how narrow these formal models are compared with appreciative theorizing. Even in the Romer–Jones model, which models the growth of the world technological frontier, we can see how conceptually (but not mathematically) simple the relationships are. The concentration on steady-state growth is particularly limiting. It has been argued that Kaldor's (1961) stylized facts of growth in fact imply steady-state growth, most notably a constant capital–output ratio (so the growth of output must equal the growth of capital) and constant factor shares. This has led to a tendency to assume that it is reasonable to assume that countries are not very far from their steady-state growth paths. However, Kaldor's stylized factors are compatible with other explanations apart from the steady-state neoclassical growth models (see Eichner and Kregel, 1975, 1977, and Brems, 1977).

There are two points to note about this. The steady-state solution implies that all sectors of the economy grow at the same rate, but the one thing we know about growth is that it involves vast structural change with some sectors (industries) initially growing fast and then declining. As Pack (1994) asks, would Japan have grown as fast as it did in the early postwar years without a deliberate policy of moving into high technology industries rather than concentrating on those industries (such as textiles) where the country had a initial (static) comparative advantage? It is doubtful if this would have occurred without government direction. Moreover, once we move away from considering the advanced countries, and turn our attention to the less developed countries we find that growth is anything but smooth (Easterly, 2001, Chapter 10; Gylfason, 1999, Chapter 1). Is it more useful to treat this as simply the result of shocks around a steady state-growth path or is the concept of the steady state itself not particularly helpful? At the end of the day, it is necessary to decide whether these models have told us anything that we did not already know. Have they provided any guidance for

policy makers that appreciative theorising has not already told us? My opinion is that they have added very little.

But, if all this were not enough, there are still even more fundamental problems with neoclassical growth theory, whether of the 'old' or the 'new' variety, to which I now turn.

SOME FUNDAMENTAL PROBLEMS WITH THE AGGREGATE PRODUCTION FUNCTION

The problems facing both the Solow and the endogenous growth models do not end here. There are other serious theoretical problems that arise with the use of the aggregate production function that have either been ignored, or indeed, forgotten. The theoretical problems posed by the Cambridge Capital Theory Controversies undermined the logical coherence of the neoclassical production function by demonstrating the insurmountable theoretical problems inherent in measuring capital (Harcourt, 1972; Cohen and Harcourt, 2003). It further showed that, comparing steady states, there is no inverse monotonic relationship between the rate of profit and the capital–labour ratio, as in the neoclassical schema. Even some neoclassical economists were disturbed by the conclusions of the controversies. Commenting on Brown's (1980) comprehensive survey of both the capital controversies and the aggregation problem, Burmeister (1980, p. 423) concluded, 'I fully agree with Brown's stated conclusion that "the neoclassical parable and its implications are generally untenable". […] Freak cases such as Samuelson's surrogate production function example are of little comfort'. He even made the revolutionary suggestion that 'for the purpose of answering many macroeconomic questions – particularly about inflation and unemployment – we should disregard the concept of a production function at the microeconomic level'. If we do this, then, of course, the concept of the production function at the macroeconomic level is also vitiated.

The second criticism concerns the possibility of summing individual production functions into a meaningful aggregate relationship that reflects the underlying technology of the economy. This is the so-called 'aggregation problem' (Fisher, 1992). There is now a large technical literature on the necessary and sufficient conditions under which micro-production functions can be aggregated to give a well-defined aggregate production function, which only are met in the most unlikeliest of circumstances (Brown, 1980; Felipe and Fisher, 2003). But the problem is intuitively very straightforward. Consider, say, the manufacturing sector. This consists of such diverse industries as (to take as random examples) the US SIC 204 Grain Mill Products and SIC 281 Industrial Organic Chemicals. Does it make any sense to combine the values of the outputs of the two industries and also their inputs and estimate a production function

that purportedly represents the underlying combined technology of these two industries? In fact, the actual position is even worse than this, as estimating an aggregate production function for, say, manufacturing, combines many such equally disparate industries, and for the total economy, even more.

Yet many studies uncritically use the aggregate production function, whether in a growth accounting context (e.g., Maddison, 1987)[12] or in econometric studies (e.g., Mankiw, Romer and Weil, 1992), using data for both the advanced and the developing countries. Fisher (2005), who over the years has done more than most to determine the technical conditions under which one can aggregate micro production functions into an aggregate production function, has summarized the conclusions of this literature as follows: 'the conditions for aggregation are so very stringent as to make the existence of aggregate production functions in real economies a non-event' (p. 490). He further argues that the conditions are such that aggregate production functions cannot even be regarded as approximations as Solow (1957), for example, regards them.

The reason why the aggregate production function continues to be so widely used is that it generally seems to work in that it gives good statistical fits with plausible parameter values. But this simply leads to the next question; given all the problems noted above, why does it work?

The reason is deceptively simple. For some time, I, together with a colleague, Jesus Felipe, have been examining this problem, following the critiques of Phelps Brown (1957), Shaikh (1974) and Simon (1979). (See Felipe and McCombie (2005a) for a summary of our arguments.) While ideally the production function should be estimated using physical data (after all, it is supposedly a technological relationship), the mere fact that heterogeneous inputs and outputs have to be aggregated means that value data must be used. From the national accounts, we can construct the following identity: $Y \equiv wL + rK$ where Y, w, L, r, K are value added, the wage rate, employment, the *ex post* rate of profit and the constant price value of the capital stock, the last calculated using gross investment by the perpetual inventory method. This identity is true for any type of production system; it holds whether or not markets are competitive; whether there are constant or increasing returns to scale; and is applicable equally to advanced or less developed countries.

Let us assume that factor shares are constant because enterprises (whether they are multinational corporations or street vendors) simply determine prices by a constant mark-up on their unit costs. These mark-ups may vary between enterprises, but we assume that the average mark-up is relatively stable. (Solow, 1958, has shown that aggregate factor shares are likely to be more stable than the underlying microeconomic shares). Expressing the identity in terms of growth rates gives the growth of value added as:

$$\hat{Y} \equiv (1-a)\hat{w} + a\hat{r} + (1-a)\hat{L} + a\hat{K} \tag{4}$$

where ˆ again denotes a proportionate growth rate. Equation (4) may be integrated to give $Y \equiv Bw^{(1-a)}r^{a}L^{(1-a)}K^{a}$, where B is a constant. If, as empirically seems to be the case, the rate of profit does not show any secular trend (as opposed to cyclical fluctuations) and the growth of the wage rate is roughly constant over the time period being considered, then equation (4) may be expressed as $Y = K^{a}(A_0e^{\lambda t}L)^{(1-a)}$ (where $\hat{w} = \lambda$ and $A_0 = B^{1/(1-\alpha)}r^{a/(1-a)}w_0$), a form identical to the Cobb–Douglas production function. It is worth re-emphasising that this derivation does *not* require that factors are paid their marginal products, either at the aggregate or the micro level and it does *not* mean that constant returns to scale exist using physical data.

I should perhaps explicitly mention that this does not deny that there is a relationship between outputs and inputs when measured in physical terms, although the exact specification may be quite complex and difficult to estimate. After all, design engineers have to have a good idea of the capacity of new plants and production processes that they design. But the statistical estimation of such engineering production functions are few and far between. What the critique does say is that it is impossible to recover estimates of technological parameters, such as the aggregate elasticity of substitution, which, of course, is a flawed theoretical concept, using value data. It explains why simulation studies, where the conditions for successful aggregation are deliberately violated (Fisher, 1971) and/or the underlying micro-production functions are not Cobb–Douglas (Nelson and Winter, 1982; Hartley, 2000; Shaikh, 2005), give very good fits to the aggregate Cobb–Douglas relationship so long as the factor shares are constant.

Hence, the fact that the Cobb–Douglas (and other more flexible production functions) give a good statistical fit to the data is no reason for inferring that an aggregate production function, or parameters such as the aggregate elasticity of substitution, actually exist even as an approximation.

We are now in a position to see why the growth accounting approach must always minimize the role played by capital accumulation in economic growth. The concept of total factor productivity assumes the existence of a well-behaved aggregate production function and that factors are paid their marginal products, notwithstanding our reservations noted above. Consequently, the growth of total factor productivity is given by $TFPG = \hat{Y} - (1 - a_t)\hat{L} - a_t\hat{K}$ where the output elasticities are equal to the relevant factor shares. Under the usual neoclassical assumptions, this is also equal to the 'dual' measure of *TFPG*, given by $(1 - a_t)\hat{w} + a_t\hat{r}$, which will approximately equal $(1 - a_t)\hat{w}$ as the rate of profit is roughly constant. But it should be noted that this also follows directly from the accounting identity, which does not require any of the usual neoclassical assumptions, including the existence of an aggregate production function. Given shares are also roughly constant, it follows that $TFPG = (1 - a_t)\hat{w} = (1 - a_t)(\hat{Y} - \hat{L})$. It is found that the share of labour in output for most countries ranges

between 0.6 and 0.75. It follows from the identity that *TFPG* will be about 60 to 75 per cent of the growth of output per capita, and slightly less of the growth of output.[13,14] We have seen above that this indeed proves empirically to be the case. The illustrative figures are lower than the percentage that Solow (1957) found in his study, but this is simply for the reason that the rate of profit, contrary to the stylized fact, was increasing over the period he considered. Solow (1988, p. 313) found the result 'startling'. It is not startling once we release ourselves from the intellectual straightjacket of the neoclassical production function and consider basic accounting identities. Felipe and McCombie (2006) show by means of some simulation exercises that the growth of total factor productivity calculated using value data can greatly differ from the 'true' rate obtained when the data are expressed in physical units.

This problem concerning the use of value data plagues all estimations of the putative neoclassical production function. As an example, let us reconsider Mankiw–Romer–Weil's (1992) model, the version of which augmented by human capital they claim supports the traditional Solow model. For expositional ease, I shall only consider the specification with physical investment, although it does not seriously affect the argument. This draws heavily on Felipe and McCombie (2005b).

Assuming only the stylized facts that the growth of output and the capital stock are equal and that factor shares are constant, the accounting identity may be written as:

$$\ln y = c + \ln w + \frac{a}{1-a}\ln r + \frac{a}{1-a}\ln s_K - \frac{a}{1-a}\ln\left[n + \delta + \frac{(1-a)\hat{w} + a\hat{r}}{1-a}\right] \quad (5)$$

Using a Cobb–Douglas production function, it can be shown that the log of the steady-state level of productivity can be expressed as:

$$\ln y = c + \ln A + \frac{a}{1-a}\ln s_K - \frac{a}{1-a}\ln(n + \delta + \lambda) \quad (6)$$

It is immediately apparent that equation (6) is close to the identity, except that both the level and the growth rates of the factor prices are implicitly treated as constant across countries in equation (6).[15] MRW, as I have already noted, find a reasonably good fit with an R^2 of 59 per cent (except for the data using just the OECD countries), but that the implied share of capital in their version of the model exceeds its factor share – although it is in the right ballpark, being about 0.6. Including human capital improves the parameter estimates and the R^2 increases to 79 per cent. 'Put simply, most international differences in living standards can be explained by differences in accumulation of both human and physical capital' (Mankiw, 1995, p. 295).

Jesus Felipe and I, however, interpret these results merely as being due to the estimation of an identity with omitted variable bias. Easterly and Levine (2001) generalize the MRW model by using dummy variables to allow ln$A$ to differ between the major regions of the world.[16] They find that 'South Asia and Sub-Saharan Africa have significantly lower productivity than the other regions (income differences that are not explained by the MRW term [i.e., $\ln s_K - \ln(n + \delta + \lambda)$]. The OECD has higher productivity than the rest of the world by a factor of 3 […] Once the productivity level is allowed to vary, the coefficient on MRW implies a capital share of .31 – which is line with most estimates from national income accounting' (Easterly and Levine, 2001, p. 190).

But we can see immediately precisely why they get this result. Because of the underlying accounting identity, it is inevitable as the dummies are closely proxying the variation of the ln$w$ term in the identity, bringing the estimated regression coefficients closer to those of the identity. If we were also to allow the growth rates of total factor productivity to vary between countries (as the neoclassical growth accounting approach shows should be the case)[17] rather than imposing a constant growth rate λ, we would end up merely estimating the identity. Consequently, it is doubtful whether this model actually tells us anything useful about why some countries are rich and others poor (Felipe and McCombie, 2005b).

We can also see why the simple AK model will give an equally good fit to the data. From the identity and the stylized facts we have $\hat{w} = \hat{Y} - \hat{L}$ and $\hat{r} \approx 0$. Consequently, substituting these into the identity given by equation (3) we obtain $\hat{Y} = \hat{K}$ or in level form $Y = AK$, where A is the constant of integration. Thus, the data are compatible with *both* approaches, but whether either helps us understand why growth rates differ is another matter.

CONCLUSIONS

This chapter has provided an assessment of the neoclassical approach to the understanding of economic growth. It will be clear that the overall conclusion will lie more with the dissenting voices mentioned in the introduction than with Mankiw (1995). It has been seen that that there are still a lot of questions concerning growth that are unanswered and that the use of the concept of the neoclassical aggregate production function is neither helpful in this respect, and nor, consequently, is the empirical work that relies on it. The recent emphasis on the role of geography, political institutions, and property rights are bringing new insights into the growth process, but there is no reason why these should be based on the production function approach. One major area of omission of the neoclassical approach is that it provides essentially closed economy models, although there are one or two exceptions. But as Tony Thirlwall has shown, one

cannot ignore the financial interrelationship between countries, caused by their trade relationships and international capital flows. Where there are under-utilized resources, or where technical change and capital accumulation are functions of the growth of output, one cannot ignore the demand side of the economy. It is here that Thirlwall has shown that the balance of payments can be a factor in constraining or limiting the maximum growth of a particular country. It is by using this approach that the understanding of why growth rates differ can be greatly advanced, rather than by trying to rescue the neoclassical story. Solow (1997, p. 232) came to the conclusion that, in another context, 'the demand-driven growth story sounds quite implausible to me under current conditions; but it is an example of the kind of question that needs to be asked'. Perhaps there is more to it than Solow thinks.

NOTES

1. Alternatively we can postulate a production function where there is capital–labour substitutability, but where the rate of interest is constant over time and so the capital–output ratio is constant. This was Harrod's preferred assumption.
2. However, as we shall see, it has been argued that the evidence does, in fact, suggest conditional convergence (the conditioning variables being the ratios of investment in physical and human capital to output).
3. It also implies a lack of either absolute or conditional convergence if all countries are in steady-state growth.
4. The data often do show sufficient variation for well-determined regression estimates of the elasticity of substitution to be obtained, *pace* Mankiw. Needless to say, this is predicated upon the existence of an aggregate production function.
5. Suppose we have a task that requires 10 workers of a certain skill level, but one of the workers is only capable of half that efficiency. This will reduce the total output by one half, and two workers capable of only half the efficiency of a skilled worker do not equal one skilled worker. Needless to say, this example is more illustrative than realistic.
6. Guyana's per capita income in 2004 in purchasing power parity terms was $4,600 while that of the US was $37,800.
7. It should be noted that $\ln A = \ln A_0 + \lambda t$.
8. I shall not, however, discuss the econometric problems that have been raised. See Felipe and McCombie (2005b).
9. Labour is subsumed in K as human capital or else must be constant for steady-state growth.
10. We assume there are no aggregation problems.
11. Conlisk (1968) showed sometime ago that increasing returns were compatible with steady-state growth provided the output elasticity of capital was less than one and the production function was a Cobb–Douglas.
12. An earlier survey of the growth accounting approach by Nadiri (1972), unlike Maddison (1987), dealt at length with the aggregation problem and the capital theory controversies, and concluded 'the aggregate production function does not have a conceptual reality of its own' (Nadiri, 1972, p. 1146). Nadiri ends his survey with a plea for more disaggregated studies, (although partly within the framework of the microeconomic production function). Temple (1999, p. 150), in his survey of the new growth theory evidence, notes that 'arguably the aggregate production function is the least satisfactory element of macroeconomics, yet many economists seem to regard this clumsy device as essential to an understanding of national income levels and growth rates' and echoes Nadiri with a plea for more studies at a more

disaggregated level. However, the problems are not avoided even at the micro level if value data are used, as we shall see.

13. We have noted that Young (1992) finds *TFPG* to be negligible for Singapore. This is equally explicable in terms of the accounting identity as the rate of profit declined sharply over the period under consideration and this offset the increase in wage rate of the efficiency adjusted labour input.
14. If labour's share is 0.60 and the growth rate of productivity is 3 per cent and of output is 4 per cent, *TFPG* will 'explain' 60 per cent and 45 per cent of these two growth rates respectively.
15. Note that under neoclassical assumptions and using the aggregate theory of factor pricing the following equations hold: $\hat{\lambda} = ((1 - a)\hat{w} + a\hat{r})/(1 - a)$ and $\ln A = ((1 - a)\ln w + a\ln r)/(1 - a)$.
16. These are OECD, East Asia, South Asia, Sub-Saharan Africa, Western Hemisphere, Middle East and North Africa, and Europe.
17. MRW are inconsistent with the standard neoclassical growth accounting results on this point, and, of course, *vice versa.*

REFERENCES

Abramovitz, M. (1952), 'Economics of Growth' in B. Haley (ed.), *A Survey of Contemporary Economics*, vol. 2, Homewood, IL: Richard D. Irwin.

Abramovitz, M. (1956), 'Resource and Output Trends in the United States Since 1870', *American Economic Review*, **46**, 5–23

Abramovitz, M. (1986), 'Catching Up, Forging Ahead and Falling Behind', *Journal of Economic History*, **46**, 385–506.

Abramovitz, M. (1993), 'The Search for the Sources of Growth: Areas of Ignorance, Old and New', *Journal of Economic History*, **53**, 217–43.

Arestis, P. and Sawyer, M.C. (2006), 'Endogenous Growth Theory: A Partial Critique', Chapter 7, this volume.

Auerbach, A.J., Hassett, K.A., and Oliner, S.D. (1994), 'Equipment Investment and Economic Growth: Comment', *Quarterly Journal of Economics*, **109**, 789–802.

Barro R. (1999), 'Notes on Growth Accounting', *Journal of Economic Growth*, **64**, 119–37.

Blomstrom, M., Lipsey, R.E. and Zejan, M. (1996), 'Is Fixed Investment the Key to Economic Growth?', *Quarterly Journal of Economics*, **111**, 269–76.

Brems, H. (1977), 'Reality and Neoclassical Theory, *Journal of Economic Literature*, **15**, 72–83.

Brown, M. (1980), 'The Measurement of Capital Aggregates. A Postreswitching Problem', in D. Usher (ed.), *The Measurement of Capital*, Chicago: University of Chicago Press.

Burmeister, E. (1980), 'Comment', in D. Usher (ed.), *The Measurement of Capital*, Chicago: University of Chicago Press.

Cesaratto, S. (1999), **'**Savings and Economic Growth in Neoclassical Theory', *Cambridge Journal of Economics*, **23**, 771–93.

Chenery, H. and Syrquin, M. (1975), *Patterns of Development, 1950–1970*, Oxford: Oxford University Press.

Cohen, A. and Harcourt, G.C. (2003), 'Whatever Happened to the Cambridge Capital Theory Controversies?', *Journal of Economic Perspectives*, **17**, 199–214.

Conlisk, J. (1968), 'Non-Constant Returns to Scale in a Neoclassical Growth Model', *International Economic Review*, **9**, 369–73.

DeLong J.B. and Summers, L.H. (1991), 'Equipment Investment and Economic Growth', *Quarterly Journal of Economics*, **106**, 445–502.
DeLong, J.B. and Summers, L.H. (1992), 'Equipment Investment and Economic Growth: How Strong is the Nexus?', *Brookings Papers on Economic Activity*, pp. 157–99.
DeLong. J.B. and Summers, L.H. (1994), 'Equipment Investment and Economic Growth: Reply', *Quarterly Journal of Economics*, **109**, 803–807.
Denison, E.F. (1967), *Why Growth Rates Differ: Postwar Experience in Nine Western Countries*, Washington, DC: Brookings Institution.
Denison, E.F. (1985), *Trends in American Economic Growth, 1929–1982*, Washington, DC: Brookings Institution.
Domar, E. (1946), 'Capital Expansion, Rate of Growth, and Employment', *Econometrica*, **14**, 137–47.
Easterly, W. (1997), 'The Ghost of Financing Gap. How the Harrod–Domar Growth Model Still Haunts Development Economics', *World Bank Policy Research Working Paper*, 1807.
Easterly, W. (2001), *The Elusive Quest for Growth. Economists' Adventures and Misadventures in the Tropics*, Cambridge, MA: MIT Press.
Easterly, W. and Levine, R. (2001), 'It's Not Factor Accumulation: Stylized Facts and Growth Models', *World Bank Economic Review*, **15**, 177–219.
Eichner, A.S. and Kregel, J.A. (1975), 'An Essay on Post-Keynesian Theory: A New Paradigm in Economics', *Journal of Economic Literature*, **13**, 1293–314.
Eichner, A.S. and Kregel, J.A. (1977), 'Reality and Neoclassical Theory: A Reply', *Journal of Economic Literature*, **15**, 83–4.
Fagerberg, J. (1994), 'Technology and International Differences in Growth Rates', *Journal of Economic Literature*, **32**, 1147–75.
Felipe, J. and Fisher, F.M. (2003), 'Aggregation in Production Functions: What Applied Economists Should Know', *Metroeconomica*, **54**, 208–62.
Felipe, J. and McCombie, J.S.L. (2001), 'Biased Technical Change, Growth Accounting, and the Conundrum of the East Asian Miracle', *Journal of Comparative Economics*, **29**, 542–65.
Felipe, J. and McCombie, J.S.L. (2005a), 'How Sound are the Foundations of the Aggregate Production Function?', *Eastern Economic Journal*, **31**, 467–88.
Felipe, J. and McCombie, J.S.L. (2005b), 'Why Are Some Countries Rich and Others Poor? A Sceptical View of Mankiw-Romer-Weil's Test of the Neoclassical Growth Model', *Metroeconomica*, **56**, 360–92.
Felipe, J. and McCombie, J.S.L. (2006), 'The Tyranny of the Identity: Growth Accounting Revisited', *International Review of Applied Economics*, **20**, 283–99.
Fine, B. (2000), 'Endogenous Growth Theory: A Critical Assessment', *Cambridge Journal of Economics*, **24**, 245–65.
Fisher, F.M. (1971), 'Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment', *The Review of Economics and Statistics*, **LIII**, 305–25.
Fisher, F.M. (1992), *Aggregation. Aggregate Production Functions and Related Topics*, ed. J. Monz, Cambridge, MA: MIT Press.
Fisher, F.M. (2005), 'Aggregate Production Functions. A Pervasive but Unpersuasive Fairytale', *Eastern Economic Journal*, **31**, 489–91.
Frankel, M. (1962), 'The Production Function in Allocation and Growth: A Synthesis', *American Economic Review*, **52**, 995–1002.
Gylfason, T. (1999), *Principles of Economic Growth*, Oxford: Oxford University Press.

Harcourt, G.C. (1972), *Some Cambridge Controversies in the Theory of Capital*, Cambridge: Cambridge University Press.
Harrod, R.F. (1939), 'An Essay in Dynamic Theory', *Economic Journal*, **49**, 14–31.
Hartley, J.E. (2000), 'Does the Solow Residual Actually Measure Changes in Technology?', *Review of Political Economy*, **12**, 27–44.
Hill, T.P. (1964), 'Growth and Investment According to International Comparisons', *Economic Journal*, **74**, 287–304.
Hsieh, C-T. (2002), 'What Explains the Industrial Revolution in East Asia? Evidence from Factor Markets', *American Economic Review*, **92**, 502–26.
Hussein, K. and Thirlwall A.P. (2000), 'The *AK* Model of 'New' Growth Theory is the Harrod–Domar Growth Equation: Investment and Growth Revisited', *Journal of Post Keynesian Economics*, **22**, 427–35.
Jones, C.I. (1995) 'R&D-Based Models of Economic Growth', *Journal of Political Economy*, **103**, 759–84.
Jones, H.G. (1975), *Introduction to Modern Theories of Growth*, London: Nelson.
Kaldor, N. (1957) 'A Model of Economic Growth', *Economic Journal*, **67**, 591–624.
Kaldor, N. (1961), 'Capital Accumulation and Economic Growth', in F.A. Lutz and D.C. Hague (eds), *The Theory of Capital*, London: Macmillan, pp. 177–222.
Kenny, C. and Williams, D. (2001), 'What Do We Know About Economic Growth? Or, Why Don't We Know Very Much?', *World Development*, **29**, 1–22.
Kremer, M. (1993), 'The O-Ring Theory of Economic Development,' *Quarterly Journal of Economics*, **108**, 551–76.
Levacic, R. and Rebman A. (1982), *Macroeconomics: An Introduction to Keynesian-Neoclassical Controversies*, Basingstoke: Macmillan (2nd edition).
Levine, R. and Renelt, D. (1992), 'A Sensitivity Analysis of Cross-Country Growth Regressions', *American Economic Review*, **82**, 942–63.
Lewis, A. (1955), *The Theory of Economic Growth*, London: Allen and Unwin.
Lucas, R.E. (1988), 'On the Mechanics of Economic Development', *Journal of Monetary Economics*, **22**, 3–42.
Lucas, R.E. (1990), 'Why Doesn't Capital Flow from Rich to Poor Countries?', *American Economic Review, Papers and Proceedings*, **80**, 92–6.
McCombie, J. S. L., Pugno, M. and Soro, B. (2002), *Productivity Growth and Economic Performance. Essays on Verdoorn's Law*, Basingstoke: Palgrave.
McCombie, J.S.L. and Thirlwall, A.P. (1994), *Economic Growth and the Balance-of-Payments Constraint*, Basingstoke: Macmillan.
McCombie, J.S.L. and Thirlwall, A.P. (eds) (2004), *Essays on Balance of Payments Constrained Growth*, London: Routledge.
Maddison, A. (1987), 'Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment', *Journal of Economic Literature*, **25**, 649–98.
Mankiw, N.G. (1995), 'The Growth of Nations', *Brookings Papers on Economic Activity*, pp. 175–310.
Mankiw, N.G. (2002), *Macroeconomics*, New York: Worth Publishers (5th edition).
Mankiw, N.G., Romer, D., and Weil, D.N. (1992), 'A Contribution to the Empirics of Economic Growth', *Quarterly Journal of Economics*, **107**, 407–37.
Nadiri, M. I. (1972), 'International Studies of Factor Inputs and Total Factor Productivity: A Brief Survey', *Review of Income and Wealth*, **18**, 129–54.
Nelson, R.R (1956), 'A Theory of the Low-level Equilibrium Trap in Underdeveloped Economies', *American Economic Review*, **46**, 894–908.
Nelson, R.R. (1973), 'Recent Exercises in Growth Accounting: New Understanding or Dead End?', *American Economic Review*, **63**, 462–8.

Nelson, R.R. (1991), 'Diffusion of Development: Post-World War II Convergence Among Advanced Industrial Nations', *American Economic Review. Papers and Proceedings*, **81**, 271–5.
Nelson, R.R. (1996), *The Sources of Economic Growth*, Cambridge, MA: Harvard University Press.
Nelson, R.R. (1998), 'The Agenda for Growth Theory: A Different Point of View', *Cambridge Journal of Economics*, **22**, 497–520.
Nelson, R.R and Pack, H. (1999), 'The Asian Miracle and Modern Growth Theory', *Economic Journal*, **109**, 416–36.
Nelson, R.R. and Winter, S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
Nelson, R.R. and Wright G. (1992), 'The Rise and Fall of American Technological Leadership: The Postwar Era in Historical Perspective', *Journal of Economic Literature*, **30**, 1931–64.
OECD (2004), *Understanding Economic Growth. Macro-level. Industry-level. Firm-Level*, Basingstoke: Palgrave Macmillan.
Pack, H. (1994), 'Endogenous Growth Theory: Intellectual Appeal and Empirical Shortcomings', *Journal of Economic Perspectives*, **8**, 55–72.
Phelps Brown E.H. (1957), 'The Meaning of the Fitted Cobb–Douglas Function', *Quarterly Journal of Economics*, **71**, 546–60.
Prescott, E.C. (1998), 'Needed: A Theory of Total Factor Productivity', *International Economic Review*, **39**, 525–52.
Pritchett, L. (1996a), 'Population Growth, Factor Accumulation and Productivity', *World Bank Policy Research Working Paper*, 1567.
Prichett, L. (1996b), 'Where Has All the Education Gone?', *World Bank Policy Research Working Paper*, 1581.
Pritchett, L. (1997), 'Divergence, Big Time', *Journal of Economic Perspectives*, **11**, 3–17.
Pritchett, L. (2000), 'The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) is Not Capital', *Journal of Economic Growth*, **4**, 361–84.
Roberts, M. (2006), 'Modelling Historical Growth: A Contribution to the Debate', Chapter 6, this volume.
Romer, D. (2001), *Advanced Macroeconomics*, Boston: McGraw-Hill (2nd edition).
Romer, P.M. (1986), 'Increasing Returns and Long-Run Growth', *Journal of Political Economy*, **94**, 1002–37.
Romer, P.M. (1987) , 'Crazy Explanations for the Productivity Slowdown', *NBER Macroeconomics Annual*, pp. 161–202.
Romer, P.M. (1990), 'Endogenous Technical Change', *Journal of Political Economy*, **94**, 71–102.
Romer, P.M. (1994), 'The Origins of Economic Growth', *Journal of Economic Perspectives*, **8**, 3–22.
Rostow, W.W. (1960), *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge: Cambridge University Press.
Rowthorn, R.E. (1999), 'Unemployment, Wage Bargaining and Capital-Labour Substitution', *Cambridge Journal of Economics*, **23**, 413–24.
Ruttan, V.W. (1997), 'Induced Innovations, Evolutionary Theory and Path Dependence: Sources of Technical Change', *Economic Journal*, **107**, 1520–29.
Sachs, J.D., McArthur, J.W., Schmidt-Traub, G., Kruk, M., Bahadhur, C., Faye, M. and McCord, G. (2004), 'Ending Africa's Poverty Trap', *Brookings Papers on Economic Activity*, Washington, DC: Brookings Institution, pp. 117–216.

Scott, M.FG. (1999), *A New View of Economic Growth*, Oxford: Oxford University Press.
Shaikh, A. (1974), 'Laws of Production and Laws of Algebra: The Humbug Production Function', *Review of Economics and Statistics*, **56**, 115–20.
Shaikh, A. (2005), 'Nonlinear Dynamics and Pseudo-Production Functions', *Eastern Economic Association*, **31**, 447–66.
Simon, H.A. (1979), 'On Parsimonious Explanations of Production Relations', *Scandinavian Journal of Economics*, **81**, 459–74.
Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, ed. E. Cannan, London: Methuen (1904, 5th edition).
Solow, R.M. (1956), 'A Contribution to the Theory of Economic Growth', *Quarterly Journal of Economics*, **70**, 65–94.
Solow, R.M. (1957), 'Technical Change and the Aggregate Production Function', *Review of Economics and Statistics*, **39**, 312–20.
Solow R.M. (1958), 'A Skeptical Note on the Constancy of Relative Shares', *American Economic Review*, **48**, 618–63.
Solow, R.M. (1988), 'Growth Theory and After', *American Economic Review*, **78**, 307–17.
Solow, R.M. (1994), 'Perspectives on Growth Theory', *Journal of Economic Perspectives*, **8**, 45–54.
Solow, R.M. (1997), 'Is There a Core of Useable Macroeconomics We Should All Believe In?', *American Economic Review*, **87**, 230–32.
Solow, R.M. (2000), *Growth Theory. An Exposition*, Oxford: Oxford University Press (2nd edition).
Swan, T.W. (1956), 'Economic Growth and Capital Accumulation', *Economic Record*, **32**, 334–61.
Temple, J. (1999), 'The New Growth Evidence', *Journal of Economic Literature*, **38**, 112–56.
Thirlwall, A.P. (1979), 'The Balance of Payments Constraint as an Explanation of International Growth Rate Differences', *Banca Nazionale del Lavoro Quarterly Review*, **128**, 45–53.
Thirlwall, A.P. (2002), *The Nature of Economic Growth. An Alternative Framework for Understanding the Performance of Nations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Todaro, M.P. and Smith, S.C. (2005), *Economic Development*, London: Longman (9th edition).
Toner, P. (1999), *Main Currents in Cumulative Causation: The Dynamics of Growth and Development*, Basingstoke: Macmillan.
Uzawa, H. (1965), 'Optimal Technical Change in an Aggregative Model of Economic Growth', *International Economic Review*, **6**, 18–31.
Wade, R. (1990), *Governing the Market*, Princeton, NJ: Princeton University Press.
Young A (1992) 'A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore' in O. Blanchard and S. Fischer (eds), *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press.
Young, A. (1995), 'The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience', *Quarterly Journal of Economics*, **110**, 641–80.

9. On the core of macroeconomic theory[1]

John Cornwall

INTRODUCTION

The 1997 edition of the *American Economic Review, Papers and Proceedings* featured a symposium in which noted macro economists presented their views on what constitutes the 'core' elements of macroeconomics, encompassing both its short-run and long-run perspectives. This core was to be presented as a set of a few propositions summarizing that part of the framework that was 'usable' in the sense that it could explain the salient historical tendencies of macroeconomic development and offer useful policy principles. Given the background of the participants, it could be expected that strong differences in views would emerge, reflecting the major differences between the cores of the two reigning schools of macroeconomics, usually referred to as New Classical Macroeconomics and New Keynesian Macroeconomics. Thus, in one sense the symposium could be seen as a continuation of a running debate within macroeconomics, with the New Classical school allegedly offering rigour and elegance of exposition as the hallmark of its approach. Of special merit was its ability to present a unified body of neoclassical economic analysis, both in its micro foundations of macroeconomics and in a growth theory that extended its short-run macro analysis, something its proponents argued the New Keynesian school had failed to achieve.

To this charge New Keynesians have responded that New Classical macroeconomics provides rigour and elegance but at too high a price. Based on the adoption of a competitive equilibrium core, it offers a micro foundation too far removed from reality to provide insight into the workings of the 'real world' micro economy, and therefore in the construction of useful macroeconomic theories and policies in either the short or long run. While we are in agreement with the charge that New Classical analysis lacks the realism needed for descriptive value and policy relevance, we offer no support in this chapter for this position. Nor do we consider it further, other than as a means of clarifying and contrasting alternative views.

With respect to the New Keynesian research programme, we recognize the solid achievements of its microeconomic theory. We see this as part of a successful and long-awaited micro research programme. By replacing the perfectly competitive neoclassical markets with imperfectly competitive Keynesian markets, it has provided a number of insightful micro economic theories of market behaviour. Indeed these theories have been widely accepted by both micro and macro economists anxious to move market analysis beyond the competitive model. However, while its willingness in short-run analysis to jettison markets that always clear is to be applauded, its unwillingness to free itself of neoclassical roots in both its short-run and long-run macro analysis has led to serious errors of commission. Furthermore, merely to free the analysis of these errors of commission would not result in an acceptable macroeconomic core. Rather the residual core would suffer from a serious lack of explanatory power, the result of what we shall refer to as errors of omission.

In this chapter an alternative macroeconomic core is developed, free of the New Keynesian model's errors of commission and omission. This is done deliberately in the form of an outline, proceeding in a step by step manner, as it requires first ridding the New Keynesian core of its errors and then building on the 'usable' residual core. The end product can be described as an extended Keynesian core, a designation that deliberately omits the prefix 'New' to confirm our alterations. It is Keynesian in its assumption that when market disequilibria occur, adjustments are induced primarily through output rather than price changes, and in its emphasis on the dominant role of aggregate demand in the determination of unemployment and growth. It is an extension of the traditional Keynesian core in describing in greater detail the macroeconomic processes at work in actual capitalist economies, both in the short run and long run, and in providing an account of the historical development of these economies.

Before undertaking these tasks, the next section takes up some 'stylized facts' of macroeconomic development. These are the salient historical tendencies we expect any macroeconomic theory to be capable of explaining. Given the rather casual disregard for the historical record in so much of current macroeconomic analysis, this rather obvious requirement is emphasized throughout the chapter.

WHAT IS TO BE EXPLAINED?

Our focus is on the short-run and longer-run macroeconomic developments of the advanced capitalist economies in a period comprising the years following World War II until the end of the century. Our sample includes all the advanced capitalist economies for which comparable data are available. Table 9.1 allows examination of the relevant macroeconomic data for eighteen economies. The total period divides into five short cycles in GDP, with common end points for

Table 9.1 *Annual average unemployment rates (U)[a], rates of consumer price inflation (p)[b] and productivity growth rates (q)[b] for 18 countries (%)*

	1960–1967			1968–1973			1974–1979			1980–89			1990–2000		
	U	*p*	*q*	*U*	*p*	*q*	*U*	*p*	*q*	*U*	*p*	*q*	*U*	*p*	*q*
Low unemployment															
Austria	2.00	3.60	4.90	1.80	5.20	5.10	1.80	6.20	2.70	3.30	3.80	1.20	3.90	2.40	1.40
Japan	1.40	5.70	8.50	1.20	7.10	7.50	1.90	9.90	2.80	2.50	2.50	2.70	3.20	1.00	1.30
Norway	2.00	3.90	3.80	1.70	6.90	2.30	1.80	8.70	2.50	2.80	8.30	1.90	4.80	2.50	2.30
Sweden	1.60	3.80	3.90	2.20	6.00	2.90	1.90	9.80	0.50	2.60	7.90	1.60	7.10	3.30	2.40
Switzerland	0.00	3.40	2.90	0.00	5.60	3.00	0.40	4.00	0.60	0.60	3.30	0.30	3.10	2.30	0.60
Unweighted Average	1.40	4.10	4.80	1.40	6.20	4.20	1.60	7.70	1.80	2.40	5.20	1.50	4.40	2.30	1.60
High unemployment															
Canada	4.80	2.40	2.60	5.40	4.60	2.50	7.20	9.20	0.60	9.40	6.50	0.90	9.30	2.20	1.30
Ireland	4.90	4.00	4.10	5.60	8.90	4.60	7.90	14.90	3.40	14.30	9.20	3.60	11.30	2.60	3.50
Italy	4.80	4.00	6.30	5.70	5.80	4.90	6.60	16.10	2.60	8.00	11.10	2.00	10.60	4.00	1.70
United States	4.90	2.00	2.60	4.60	5.00	1.00	6.80	8.50	0.50	7.30	5.50	1.20	5.60	3.00	1.90
Unweighted Average	4.90	3.10	3.90	5.30	6.10	3.30	7.10	12.20	1.80	9.80	8.10	1.90	9.20	3.00	2.10

Low-high unemployment															
Australia	2.20	2.20	2.70	2.00	5.60	2.20	5.10	12.20	1.80	7.50	8.40	1.00	8.40	2.70	1.80
Belgium	2.00	2.80	3.90	2.50	4.90	4.90	7.10	8.40	2.40	9.80	4.90	2.20	8.50	2.20	1.60
Denmark	1.60	6.20	3.00	1.00	6.30	3.00	6.10	10.80	1.20	8.10	6.90	0.70	7.10	2.20	1.90
Finland	1.60	5.60	4.00	2.60	5.80	5.80	5.10	12.60	1.70	5.40	7.10	2.60	11.70	2.30	2.60
France	1.60	3.60	4.90	2.60	6.10	4.30	4.50	10.70	2.40	8.80	7.30	2.20	11.10	1.90	1.30
Germany	0.60	2.70	4.10	1.00	4.60	4.00	3.20	4.60	2.70	5.80	2.90	1.50	7.70	2.50	1.60
Netherlands	1.00	3.60	3.70	1.50	6.90	4.40	5.40	7.20	2.10	7.90	2.80	–0.30	5.40	2.50	0.80
New Zealand	0.10	3.30	1.20	0.30	7.40	2.70	0.80	13.80	–1.60	4.60	11.80	0.50	7.80	2.10	0.80
United Kingdom	2.70	3.60	2.70	3.30	7.50	3.20	4.70	15.60	1.20	9.80	7.40	1.90	8.00	3.60	1.80
Unweighted average	1.50	3.70	3.40	1.90	6.10	3.80	4.70	10.70	1.50	7.50	6.60	1.40	8.40	2.40	1.60
Overall Average	2.20	3.70	3.90	2.50	6.10	3.80	4.40	10.20	1.70	6.60	6.50	1.50	7.50	2.50	1.70

Notes:

[a] OECD *Historical Statistics 1970–2000* and *1960–1997* (Paris: OECD), Table 2.19, Standardized Unemployment Rates. Data for 1960–64 are from the OECD–CEP Data Set (*Working Paper 118*, London: Centre for Economic Performance). For Austria, Denmark and Switzerland, and for New Zealand prior to 1974, standardized rates are not available; unemployment as a percentage of the total labour force is used instead.

[b] OECD *Historical Statistics 1970–2000* and *1960–1997* (Paris: OECD), Table 7.10 Consumer Price Inflation and Table 3.7 Real GDP per person employed.

each economy indicated at the top row of the table. Rates of unemployment, inflation and growth of labour productivity are listed for each country and each short-run GDP cycle.[2] Considering the entire post World War II period, the economies fall into three groups: the 'low unemployment economies' with full employment in the Golden Age (1960–73) and in most of the subsequent short-run cycles; the 'high unemployment economies' with high unemployment rates in excess of full employment both during and after the Golden Age; and the 'low-high unemployment economies' with full employment during the Golden Age followed by high unemployment subsequently.

The total post war period also divides naturally into two longer episodes, what we have above called the Golden Age from the years following World War II until the mid-1970s, and what we will designate the Age of Decline from the mid-1970s to the end of the century. The first episode includes the first two short-run GDP cycles and the second includes the last three successive within-episode cycles. These episodes are distinguished by manifestations of markedly different macroeconomic performance. Using 3 per cent as the full employment rate of unemployment, in the Golden Age episode all but four of the eighteen countries in the table experienced full employment. In the Age of Decline only three economies came close to achieving full employment. In general, rates of growth of labour productivity traced out a rapid growth–slow growth pattern from the Golden Age episode to the Age of Decline. This marked historical shift in macro performance from the first half to the second half of the entire period has prompted some macro economists to consider a third analytical category worthy of separate analysis when modelling the postwar era, the medium run, (Blanchard 1997a, 1997b; Solow, 1997, 2000; Cornwall and Cornwall, chs. 10 and 11, 2001). Our long run can then be seen as a succession of two medium runs, or what we call episodes of macro performance.

We will argue that an acceptable model of growth and development for the period covered in this study can only be formulated within a short-run, medium-run and long-run format. Furthermore, in spite of the cross-country differences in macro performance as just cited, this remains true with respect to modelling both the performance of individual economies and that of the eighteen economies as a group. We also will argue that New Keynesian macroeconomics is ill-suited to this important task, both for its failures to incorporate a medium run as a part of its core model and to explain the shifts in performance from one historical episode to the next. With the macroeconomic record firmly in mind, we turn to a brief description of the New Keynesian core, followed by two sections in which the 'fit' between the analytical features of our core and the stylized historical facts are discussed.

THE CORE OF THE NEW KEYNESIAN MACROECONOMICS

The Dichotomies

Following a long-established tradition in economic theory, the core of New Keynesian macroeconomics adopts a dichotomous format, limiting time periods in the analysis to the short and long run. The medium run has been a recent 'discovery'. Two similar dichotomies have been used. Within each classification scheme the presence of mutually exclusive structural characteristics of the economy are specified, which distinguish the short run from the long run period. These distinguish between the short run and long run either in terms of differences in macroeconomic processes at work or in terms of differences in market structures. Specifically, in the first classification, demand forces are assumed to dominate macro movements, generating short cycles or convergence movements to short-run equilibria, with forces on the supply side dominating long-run developments, i.e., growth rates. With respect to the dichotomy based on market structures, in the short run Keynesian markets are assumed to prevail. Market disequilibria induce adjustments in output rather than prices, and involuntary unemployment is a possibility. In the long run neoclassical markets dominate, with rapid or even instantaneous market clearing, and full employment conditions prevail. While there is some overlapping of the two dichotomies, serious problems arise whichever is adopted, a matter discussed in the next section.

The Short Run

There are two short-run versions of the core from which to choose . The well-known textbook form is distinguished by the key role played by the vertical Phillips curve. In this version, Keynesian models of imperfectly competitive labour and product markets implicitly serve as the micro support system for the demand side of a short-run macro model of the IS–LM variety. Short-run movements of output are treated as the result of monetary shocks generating temporary changes in output and employment, followed by a steady convergence back to a unique (exogenous) NAIRU equilibrium. Individual markets are Keynesian to the extent that they do not necessarily clear, as market imperfections slow the speed of adjustment to shocks. But in the aggregate, either demand adjusts passively and inevitably to the exogenous 'attractor' or else the monetary authorities can always move things along to some exogenous output or unemployment target by means of minor adjustments in interest rates or the money supply.[3] In either event, even while New Keynesians assume markets are Keynesian in the short run and designate their analysis a form of Keynesianism, aggregate demand plays no independent role even in the determination of short-

run equilibrium. Successive short-run movements then trace out a succession of adjustments to supply determined NAIRU equilibria. This most un-Keynesian property of the core reveals an unwillingness of most New Keynesians to free their short-run analyses of their neoclassical roots.

The alternative short-run formulation also incorporates Keynesian markets but does not retain the unique equilibrium tendencies of NAIRU analysis (Blinder,1988; Solow, 1997). Rather some form of Keynesian income–expenditure mechanism is chosen to model short-run dynamics. Unlike the first version of the short-run model, this formulation can be described as consistently Keynesian, as it assumes demand determined short-run equilibria rather than a unique NAIRU equilibrium.[4] Movements of unemployment rates from one short-run period to the next over the total period covered in Table 9.1 are interpreted as a succession of short-run damped cycles in GDP. Each traces the interactions of a Keynesian income expenditure model in a market system displaying wage and price rigidities. Given its historical record as a short run predictive device and the growing dissatisfaction of macro economists with NAIRU analysis as an analytical tool, we incorporate this second more Keynesian short-run formulation in our reconstruction efforts below.

The Long Run

New Keynesians have shown little interest in providing theoretical support for the adopted long run part of their core model or in integrating it with the short-run macro analysis. There are two types of New Keynesian formulations of the long run, each formulated as part of a long run–short run format. One is to simply fit a smoothly evolving trend line to an historical output or unemployment series, and to designate this line as the long-run equilibrium path of the economy. In effect, this mechanical approach assumes the moving equilibrium of the system can always be derived from the past history of the economy, whatever path this has taken. Although quite popular in the textbooks for over a decade (Mankiw, 1994, figure 5.1; CEPR, 1995, fn. 20), no theoretical or empirical evidence is provided to support this assumption. This formulation need not be considered further.

A second formulation models long-run growth as a supply-side process summarized by an aggregate production function. Its elements are the rates of growth of capital, labour and technical progress, the last being a residual component measuring the rate of growth in the efficiency of one or more of the factor inputs. The most popular version of New Keynesian long-run analysis is to attach a neoclassical production function, broadly defined to include new growth theory production functions, to the core short-run Keynesian model (Scott, 1989; Solow, 2000). In this way, the long-run equilibrium of the economy becomes the full employment rate of output growth, as aggregate demand is assumed to

automatically adjust to aggregate supply. Introducing shocks simply leads to the occurrence of mild, shock-induced deviations around a steady state full employment growth path.

ERRORS OF COMMISSION

What must be emphasized is that attempts to construct a New Keynesian core from a combination of any one of the short-run and the long-run formulations just discussed brings out a serious internal inconsistency in its research programme. In both dichotomies, the problem arises from an insistence on drawing an analytical distinction between the short run and long run, defined by the presence or absence of mutually exclusive structural characteristics. Consider the short-run Keynesian and long-run neoclassical markets dichotomy as an example . Assume the long-run growth path is modelled by a neoclassical production function, with the growth of supply being exogenous and independent of aggregate demand. In assuming the economy moves along a full employment growth path, an internal inconsistency arises. Any Keynesian short run must eventually become part of a neoclassical long run. Yet there is no explanation within the core of when and how the economy undergoes the radical changes in industrial structure required if competitive forces are to eventually dominate Keynesian market influences. To put it another way, the long run cannot be a simple extension of the short run when the classification criterion assumes different structural features are introduced (or deleted) in the long run. This error-of-commission criticism holds equally well with the dichotomy formulated in terms of the macroeconomic processes at work. In this case, there is no explanation within the core of when and how the economy undergoes the radical shift in its structural framework required to transform the economy from a demand driven to a supply driven one

The question arises why such fundamental flaws in construction of its core has remained uncorrected by New Keynesian macroeconomists. One clue is provided by a recognition that any such inconsistencies are of little consequence if the short-run performance of the economy is assumed to have little effect on its long run performance. Indeed, as pointed out in the third section above, this assumption plays an important role in New Keynesian analysis. For example, short-run movements are assumed to reflect little more than mild demand induced deviations from the supply determined steady state.

However, 'resolving' this internal inconsistency in this manner merely replaces one unacceptable assumption in the New Keynesian core by another. In assuming that in the long run the economy moves along a full employment path, the New Keynesian model fails as a descriptive device; it introduces an empirical inconsistency. It assumes that the historical record can be modelled as one of

more or less continuous full employment, perhaps occasionally and briefly interrupted by shock-induced deviations from full employment. But in Table 9.1 the actual unemployment record of the advanced capitalist economies indicates a lengthy episode of high unemployment rates in a large majority of the economies.[5] While mild increases in unemployment rates within an episode might be caused by shocks, a medium-run period of persistently high unemployment cannot be so easily brushed aside. Periods of high unemployment have been too widely dispersed across economies as well as too lengthy to simply ignore. Instead, we shall argue below that an episode such as the Age of Decline indicates a lengthy period of inadequate aggregate demand and high involuntary unemployment. Errors of commission such as these indicate a major rebuilding effort is required in order to construct a usable macroeconomic core.

TRADITIONAL KEYNESIAN CORE

The first step in formulating our alternative core is to remove the features of the New Keynesian core leading to these inconsistencies. Ridding the model of its internal inconsistencies is in one sense a simple matter: model the economy as one composed of either neoclassical or Keynesian markets and allow demand and supply forces to affect performance whatever the time frame. New Classical macro economists have chosen to model markets as neoclassical and to retain the assumption that only supply forces affect long run performance. Here the decision is both to model markets as Keynesian and allow both demand and supply forces to affect outcomes in the short, medium and long runs. This eliminates the internal and empirical inconsistencies in the analysis. Unfortunately, in eliminating the inconsistencies of the New Keynesian core, important questions arise that cannot be answered within the residual New Keynesian core. These result from what we referred to in the Introduction as errors of omission.

To see this, note that in the adoption of the assumption of Keynesian markets for all time frames, the altered macro core on the demand side bears a strong resemblance to the core of Keynesian macroeconomics of the *General Theory*. In traditional Keynesian macro theory, aggregate demand drives the system, with the level of total private demand determined through the multiplier by an exogenous force, the 'animal spirits' of business community. If total demand by the private sector is not strong enough to deliver full employment, it is assumed that another exogenous force, the fiscal authorities, can provide the necessary additional stimulus.[6] This raises the question of how could it happen that for an extended historical period the authorities on average were unwilling to provide the aggregate demand stimulus required to prevent continued high if not mass unemployment? Furthermore, if markets are Keynesian whatever the

time period and high rates of unemployment are always a possibility, how could it happen that during the Golden Age the authorities were able to sustain a full employment policy stance without experiencing unacceptable side effects, e.g., accelerating rates of inflation or sustained balance of payments deficits? The traditional Keynesian core provides no answers to either question; evidently something important has been omitted.

This brings the analysis to the next step in formulating our extended Keynesian core. To explain these events, and more particularly to explain medium-run performance, two extensions of the traditional Keynesian core are required. One is to endogenize its exogenous determinants of aggregate demand policy and the other is to endogenize its exogenous determinants of aggregate supply.[7] In this way aggregate demand forces influences aggregate supply in both the short and long run.

EXTENDING THE TRADITIONAL KEYNESIAN CORE – ENDOGENIZING AGGREGATE DEMAND POLICY

In this section we present a political economy theory of aggregate demand policy as the first extension of the traditional Keynesian core. In so doing, aggregate demand policy outcomes and therefore levels of aggregate demand and unemployment rates are endogenously determined. Thus, rather than treating aggregate demand policies as given from outside the model, we deepen the analysis by investigating *why* the authorities allow extended periods of insufficient aggregate demand, while in other periods the authorities actively intervene in order to achieve the full employment goal.

To answer these questions, we model the *dominant* macroeconomic policy response of the authorities within any medium run episode as the outcome of an interaction between the supply of and demand for full employment policies.[8] The strength of demand for full employment policies is determined by the distribution of political and economic power among organized interest groups. The policies supplied by the authorities depend upon whether or not there are constraints on aggregate demand policies limiting their policy options. For example, if full employment levels of aggregate demand are seen to generate unacceptable inflation or unsustainable payments deficits, or if there are laws that limit budget deficits, less than full employment levels of aggregate demand policies will be adopted. In the absence of such constraints on aggregate demand, the political economy theory of aggregate demand policy assumes the fiscal policy authorities will provide the necessary stimulus to achieve full employment levels of aggregate demand, provided they are demanded.

The party control theory of economic policy is the most prominent of the models focussing on the demand side, offering a political explanation of fiscal

policy choice and differences in unemployment rates across countries in terms of the relative strength of right-wing and left-wing political parties (e.g. Kalecki, 1971; Hibbs, 1987; Alesina *et al.*, 1997). This is assumed to depend upon the distribution of power between capital and labour. According to this theory, labour is more willing than capital to trade price stability for lower unemployment; this preference is registered at the ballot box through its choice of political parties. From one episode to the next, differences in aggregate demand policies and unemployment are traced to shifts in political power within an economy.

However, the impact of the distribution of political power on unemployment rates can only be part of the story, accounting only for the strength of *demand* for expansionary policies. For example, even a strongly pro-labour government must consider the costs of supplying a full employment policy, the most obvious being unacceptable inflation costs or unsustainable payments deficits or both. Pushing the chain of causation even deeper, when such undesirable effects of full employment policies are evident, they are then traced to unfavourable institutions such as an adversarial industrial relations system or to an absence of institutions promoting strong export growth. An absence of such constraints on aggregate demand can be traced to a favourable set of the same institutions.[9]

A MEDIUM- AND LONG-RUN KEYNESIAN MODEL OF AGGREGATE DEMAND AND UNEMPLOYMENT

The political economy theory of aggregate demand policy gives a deeper understanding of differences in unemployment performance shown in Table 9.1. The Golden Age was an episode characterized by the absence of constraints on aggregate demand polices in most of the world's capitalist economies and the presence of an historically strong power position for labour relative to capital. As a result both the willingness of the authorities to supply full employment policies and strong political demands for full employment policies were present, leading to low unemployment rates in the Golden Age. In the episode that followed the Golden Age, most of the economies found themselves faced with institutional constraints on expansionary aggregate demand policies. Full employment was seen as incompatible with politically acceptable rates of inflation or sustainable balance of payments positions or both. Among the new constraints was the introduction of an international monetary system which deregulated international capital flows and the increased flexibility of exchange rates. These contributed to a proliferation of restrictive policies and the rise of unemployment rates most everywhere. The Age of Decline was also an episode in which labour's power was weak compared to its relatively strong position in the Golden Age and this led to weak demands for high unemployment policies.

When these medium-run episodes are considered in sequence, they form an historical period that can be modelled as a long-run theory of aggregate demand and unemployment. Each of the two episodes is a medium run in the economy's long-run development, beginning and ending with a marked change in unemployment rates. Each episode is therefore characterized by a given set of key institutions and distribution of power. A sustained radical alteration in performance signals the arrival of a new episode, characterized by new institutions or a new power distribution (or both), and a major shift in the dominant policy stance of the authorities. For example, the institutional shift from a cooperative industrial relations system to an adversarial one or a radical change in the international monetary system can lead to incompatibility of full employment with acceptable inflation rates or sustainable payments positions, or both, and restrictive policies that would end a medium-run boom episode. A radical shift in the distribution of power is also a possible source of radical shifts in macroeconomic policy and performance.[10] While we have concentrated on the eighteen capitalist economies as a group, as stated earlier we maintain that an acceptable explanation of unemployment trends in most of the individual economies of Table 9.1 can be modelled by our model of long-run aggregate demand and unemployment.

EXTENDING THE TRADITIONAL KEYNESIAN CORE BY ENDOGENIZING AGGREGATE SUPPLY

Recall that a key feature of the New Keynesian core is a dichotomy format in which different forces dominate the demand and supply side movements of the total economy. Our second extension of the traditional Keynesian core emphasizes the *empirical* incorrectness of this assumption. We argue that in the real world the alleged exogenous determinants of the supply of output cannot be considered as completely exogenous, but are to a large extent endogenous (León-Ledesma and Thirlwall, *op.cit.*). Thus, in standard production function analysis, rates of growth of capital and labour inputs and technical progress are usually assumed to be exogenous (and the immediate) determinants of rates of growth of aggregate output and labour productivity.[11] The endogenizing of what are incorrectly assumed to be exogenous determinants of a supply side growth process begins by treating each to an important degree as a function of the actual performance of the economy. Specifically, we assume each to be importantly influenced by aggregate demand. Indeed we maintain that causation can run from demand to supply whatever the time frame of the analysis.

To emphasize this essential point, consider an economy faced with no constraints on aggregate demand policies and therefore operating at full employment and a standard or normal rate of utilization of the capital stock. Initially, consider

the case in which inputs in the production process are fixed in the following sense. The capital stock measured in physical units and the age and gender distribution of a given population are fixed, and the existing technology is defined by a given set of blueprints of available production techniques. Allow a sustained small-percentage increase in the rate of growth of real aggregate demand and consider the following selection of responses from a real world catalogue of supply responses. According to Okun's law, the higher rates of growth of real demand will induce responses on the supply side in the form of higher rates of labour-force participation, overtime shifts, and a shift from part-time to full-time work by some of the already employed. This demand-induced expansion of employment will be accompanied by higher utilization rates for capital and a growth in unfilled orders in industries that produce to order rather than for inventory. In addition, the induced higher rate of growth of output will lead to higher rates of productivity growth because of 'learning-by-doing' effects. The point to be stressed is that even when it is assumed that the stocks of capital, labour and technology are fixed, the elasticity of supply with respect to aggregate demand will be positive. Demand forces influence the supply of output and productivity, even in the short run.

Next assume a continuation of the period of strong aggregate demand growth, such as the Golden Age. The cumulative effect of the increased capital utilization rates and tighter labour markets can be expected to have a pronounced and lasting impact on factor supplies and technical progress. For example, a sustained higher rate of growth of aggregate demand will pull labour out of low income elasticity–high productivity growth sectors, e.g., agriculture, to satisfy the now more rapid rates of growth of demand for labour in the industrial and service sectors. If this source of 'surplus labour' proves insufficient, the growing number of job vacancies will be filled by importing more labour from abroad. In this scenario, labour supply is endogenous and not part of the 'givens'.

The sustained additional stimulus to the rate of growth of aggregate demand will also have a positive sustained effect on the rate of growth of capital and the rate of introduction of new technologies, much of which will be technology transfer from the industrial leader(s). As Domar pointed out half a century ago, investment generates output through the multiplier but also leads to an increase in maximum output and labour productivity by increasing the capital stock. By generating a more rapid rate of growth of investment, capital, and technical progress, higher rates of growth of aggregate demand also contribute to higher rates of growth of maximum output and productivity.

It must be stressed that other influences than aggregate demand pressures affect the supply side categories and have historically contributed to the differences in rates of growth of output and, as shown in Table 9.1, differences in rates of growth of productivity both across countries and over time. But the important implication of arguments presented in this section is that strong, sus-

tained demand pressures were a necessary condition not only for the low unemployment rates, but also for the high rates of growth of productivity and output during the golden age, primarily because of their positive impact on the rates of growth of the labour force, capital, and technical progress. Similarly, stagnant demand conditions have been a key factor accounting for the poor macroeconomic performance of the last two and a half decades.

A KEYNESIAN MODEL OF UNEMPLOYMENT AND GROWTH

Two points emerge from the discussion in the previous section. First, endogenizing the traditional supply side categories provides a key element in the construction of a medium- and long-run theory of growth of output and labour productivity to complement the medium- and long-run Keynesian model of aggregate demand and unemployment presented above. Second, the long-run performance of unemployment, output and productivity can be modelled as an interaction of aggregate demand and aggregate supply, giving rise to a two-stage recursive process. As outlined above, certain key institutions and the distribution of power determine the dominant aggregate demand policy and therefore the overall strength of aggregate demand pressures and average unemployment rates in any episode. With the level of aggregate demand and the unemployment rate determined, and with due account taken of exogenous influences on performance,the average rates of growth of output and productivity in each episode are then determined. In this manner the extended Keynesian core incorporates a mechanism in which changes in the long-run growth of aggregate demand always generate changes in the long-run growth of aggregate supply in the same direction.[12] This contrasts with the dominance of aggregate supply, central to New Keynesian growth theory as well as to and New Classical growth theory.

Our theory can also be expressed in a manner that stresses its political economy character. This emphasizes the direct effects of aggregate demand policy on aggregate demand and unemployment, as well as their impact on aggregate supply. More specifically, we focus on the way in which macroeconomic performance in a modern economy can be modelled as an interaction between the economy's performance and the authorities' policy responses, a political process, and the impact of government policy intervention on the performance of the economy, an economic process. We argue that in this way our extended Keynesian macro core gains appreciably in descriptive value and policy relevance.

Absent from this outline is any consideration of possible linkages between the episodes. Given space limitations, the treatment of possible linkages between

the episodes can only be touched upon. With respect to the post World War II long-run period, there are three possibilities. The first is that there were no major causal linkages and the Golden Age episode was brought to an end by a series of shocks and our long-run model qualifies as an example of complexity theory. The second possibility is that the Age of Decline episode emerged as the result of the accumulation of unfavourable developments during the Golden Age. In this case our long-run model is an example of an evolutionary-Keynesian growth process. The third is a combination of the first two. We have argued elsewhere in favour of the third interpretation, with institutional and power developments during the Golden Age episode causally linking it with the Age of Decline. This linkage was reinforced by a series of shocks.[13]

CONCLUSIONS

The aim of this chapter has been to describe, in broad terms, the macroeconomic processes at work in actual capitalist economies and to provide an account of the historical development of these economies. This involved outlining, in a step-by-step manner, a macroeconomic core free of the shortcomings of New Keynesian macroeconomics, i.e. what we designated as its errors of commission and omission. Ultimately these shortcomings were traced to its inability to free itself from its neoclassical roots, both in its short-run and long-run analysis. Thus, while New Keynesians assume Keynesian markets in the short run, their short-run NAIRU equilibrium is determined by exogenous forces independent of aggregate demand. Their long-run analysis assumes supply determined outcomes in neoclassical markets.

Having freed the New Keynesian core of its errors of commission, the residual core, which is also the traditional Keynesian core, was then extended to take account of errors of omission. The extended Keynesian core was characterized as a demand determined model of unemployment and growth within a political economy framework. This description emphasized our endogenous treatment of government policy and its impact on macroeconomic performance.

To arguments that the proposed alterations of the mainstream macro core are extreme, our response is that the shortcomings of mainstream New Keynesian macroeconomics are sufficiently serious to justify major alterations. For example, we argued that the dichotomy formats adopted by New Keynesians are unacceptable. These can only serve to maintain the fiction that capitalist economies are self-regulating in some vague long-run sense. This flies in the face of the historical record. Table 9.1 depicts a macroeconomic record of a prolonged period of mass unemployment and politically unacceptable rates of inflation in a dozen and a half developed capitalist economies. Earlier data beginning in the 1920s, reveal the same absence of self-regulating tendencies.

Indeed, important examples of contrasts between markets structures and macroeconomic processes at work in real economies and the New Keynesian choice of modelling methods leads us to challenge of another basic feature of the New Keynesian core, the central role of equilibrium analysis in mainstream macroeconomics. In the third section, we chose to reject vertical Phillips curve analysis as a short-run modelling device, choosing a more Keynesian core to model the short run. We would only add that discarding short-run unique equilibrium analysis should not be a cause for concern. From a policy viewpoint, the important issue is whether or not shocks to an ongoing macroeconomic process have disturbance-amplifying tendencies. Exploring such possibilities is better understood by modelling short-run movements as outcomes of a Keynesian income generating mechanism that incorporates the now widely recognized wage and price inflexibilities.

The historical record also indicates the inappropriateness of modelling the long run as some kind of equilibrium process. Nevertheless New Keynesians (and Neoclassical) growth theorists are insistent that the long run be modelled as a steady state process of balanced growth, in which sectoral and institutional changes and the redistribution of power can be ignored even though the long run may extend in historical time over several decades (Blanchard, 1997b). The historical record indicates clearly the inappropriateness of their theories for modelling the 'stylized facts' of long-run macroeconomic development. Our long-run model should be seen as one in which structural change and transformation emerge as a natural elements of the dynamics of economic growth.

NOTES

1. Financial support for this research has been provided by the Office of the Vice-President Academic & Provost, the Dean of the Faculty of Science, and the Department of Economics, all of Dalhousie University. The author wishes to thank Wendy Cornwall, Mark Setterfield and John Smithins for their helpful comments and criticisms.
2. Comparable data are not available for all eighteen countries prior to 1960. The year 2000 marks the end of the last complete cycle. The common end points for the periods shown in the table are those published in *OECD Historical Statistics*.
3. The standard textbook version of the New Keynesian short-run model has been updated in the 'new consensus' theory of inflation targeting. For a critique see Arestis and Sawyer (2003) and Setterfield (2003).
4. See Cornwall and Cornwall (2001, Ch.3) for a critique of vertical Phillips curve analysis.
5. The historical record also reveals another episode of high, widespread rates of unemployment, the Great Depression of the 1930s. See table 2.1 in Cornwall and Cornwall (2001). Figures are based on data for 16 of the 18 OECD economics of Table 9.1 of this chapter; data are not available for Ireland and New Zealand for years prior to the Golden Age.
6. For an argument supporting the policy effectiveness of fiscal policy see Arestis and Sawyer (2003).
7. Economists familiar with the work of Michael Kalecki (1971) and Tony Thirlwall (1979, 2001; León-Ledesma and Thirlwall, 2002) will recognize the similarities between these extensions and their important work.

8. See Gordon, (1975), for an early explanation of policy outcomes stressing both demand and supply influences.
9. As the main aim of this chapter is to outline the core of an alternative Keynesian framework, a fuller exposition of these deeper causes is not essential. A fully specified version of this approach is tested econometrically in W. Cornwall (1999).
10. Kalecki's (1971) famous model of the political business cycle offers an alternative explanation of the interruption of an episode of strong aggregate demand, i.e. capital's ability to force governments to enact high unemployment policies.
11. The notable exception here is New Growth Theory's treatment of technical progress, by which technical change is endogenously determined on the supply side in several models . However, as the text makes clear, the focus in the extended Keynesian core is on technical progress being importantly determined on the demand side.
12. An econometric analysis of the model of long-run performance of unemployment and productivity growth rates covering the eighteen countries listed in Table 9.1 during the post World War II period can be found in Cornwall and Cornwall (2001, chs. 5, 10, 11). Mark Setterfield's contribution to this volume adopts a steady state approach to show how equality of demand and supply growth can be achieved in a demand-led growth framework.
13. See Cornwall and Cornwall, *ibid.* chs. 6 and 11.

REFERENCES

Arestis, P. and M. Sawyer (2003), 'Reinventing Fiscal Policy', *Journal of Post Keynesian Economics*, Fall.

Alesina, A. (1997), *Political Cycles and the Macroeconomy*, Cambridge, MA: MIT Press.

Blanchard, O. (1997a), 'Is There a Core of Usable Macroeconomics?', *American Economic Review Papers and Proceedings*, **2**.

Blanchard, O. (1997b), 'The Medium Run', *Brookings Papers on Economic Activity*, *No. 2.*

Blinder, A. (1998), 'The Fall and Rise of Keynesian Economics', *The Economic Record*, December.

CEPR (1995), *Unemployment Choices for Europe*, London: Centre for Economic Policy Research.

Cornwall, J. and W. Cornwall (2001), *Capitalist Development in the Twentieth Century: An Evolutionary-Keynesian Analysis*, Cambridge: Cambridge University Press.

Cornwall, W. (1999), 'The Institutional Determinants of Unemployment', in M. Setterfield (ed.), *Growth, Employment and Inflation: Essays in Honour of John Cornwall*, London: Macmillan.

Gordon, R.J. (1975), 'The Demand For and Supply of Inflation', *Journal of Law and Economics*, **18**.

Hibbs, D. (1987), 'Political Parties and Macroeconomic Policy' in *The Political Economy of Industrial Democracies*, Cambridge, MA: Harvard University Press.

Kalecki, M. (1971), 'Political Aspects of Full Employment' in *Selected Essays on the Dynamics of the Capitalist Economy*, Cambridge: Cambridge University Press.

León-Ledesma, M. and A.P. Thirlwall (2002), 'The Endogeneity of the Natural Rate of Growth', *Cambridge Journal of Economics*, **26**.

Mankiw, N.G. (1994), *Macroeconomics*, 2nd Edition, New York: Worth Publishers.

Scott, M. (1989), *A New View of Economic Growth*, Oxford: Clarendon Press.

Setterfield, M. (2003), 'Supply and Demand in the Theory of Long-run Growth: Introduction to a Symposium on Demand-led Growth', *Review of Political Economy*, **1**.

Setterfield, M. (2005), 'Central Bank Behaviour and the Stability of Macroeconomic Equilibrium', in P. Arestis, M. Baddeley and J. McCombie (eds), *The New Monetary Policy: Implications and Relevance*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 23–49.

Solow, R. (1997), 'Is There a Core of Usable Macroeconomics that We Should all Believe In?', *American Economic Review Papers and Proceeding*, **2**.

Solow, R. (2000), 'Toward a Macroeconomics of the Medium Run', *Journal of Economic Perspectives*, Winter.

Thirlwall, A. (1979), 'The Balance of Payments Constraints as an Explanation of International Growth Rate Differences', *Banco Nazionale del Lavoro Quarterly Review*, March.

Thirlwall, A. (2001), 'The Relation between the Warranted Rate of Growth, the Natural Rate of Growth and the Balance of Payments Equilibrium Growth Rate', *Journal of Post Keynesian Economics*, **24**, 81–8.

10. The crisis of the stability pact and a proposal

Roberto Tamborini and Ferdinando Targetti

INTRODUCTION

At the beginning of this century, three events seemed to be giving robust impetus to the process of European unification: the birth of a single currency, enlargement to the East of the Union, and the creation of the Convention, a sort of European constituent assembly. But then, and in a partly unexpected manner, three contrary circumstances simultaneously arose: on the foreign policy front, the split caused by the war in Iraq; on the institutional front, the failure by the Intergovernmental Conference to ratify the 'Giscard-Amato draft'; and on the economic–political one, the majority decision to suspend the validity of a treaty (the Stability and Growth Pact, SGP) taken by Ecofin (the Council of the Union's financial ministers), contrary to the Commission's recommendation. These obstacles seem so difficult to overcome that Michel Rocard (*Le Monde*, 28 November, 2004) has gone so far as to say that 'Europe has lost its soul along the way'; and perhaps not only its soul but its body as well if, as Romano Prodi has said, Europe risks disappearing from geographical maps.

The SGP is perhaps the most acute manifestation of the Union's institutional malaise, with implications that extend beyond the sphere of economic policy and affect the very foundations of the 'European common house'. The Treaty of Maastricht, supplemented by the SGP, in fact, prescribes that if a member country's budget deficit exceeds 3% of GDP (purged of the estimated effects of the economic cycle), the Commission must first warn that country to respect the parameters, and then apply pecuniary sanctions if its deficit persists for more than two years. In 2001, with Portugal, the procedure was followed to the letter; but in November 2004, when France and Germany were also found in breach of the 3% criterion, contrary to the Commission's recommendations, Ecofin decided by majority vote that the procedures laid down by the SGP should be suspended.[1]

Our intention in this chapter is to discuss the causes of the crisis of the SGP – causes which are complex because they comprise interweaving economic, legal and political-institutional factors – and then to put forward a proposal, both Europeist and realistic, for resolving the current impasse.

THE ECONOMIC DEBATE

On strictly economic grounds, one must start from the premise that the SGP was born inherently weak and contradictory. Simplifying to the extreme, the SGP has been conceived as follows. Confidence in the euro will be created and maintained if the European Central Bank (ECB) fulfils its sole institutional purpose of controlling inflation, and if the aggregate public budget net of cyclical fluctuations remains in balance. This objective requires the imposition of constraints on national public budgets for three alleged reasons: (1) The creation of a single monetary and financial market means that the deficit of a single country will absorb common financial resources, thereby increasing interest rates for all and causing collective damage in the form of reduced private investments (crowding out); (2) European governments have always tended to accumulate deficits in periods of recession, but not to correct them in periods of recovery, so that in the twenty years prior to the SGP the aggregate debt/GDP ratio of the euro-countries more than doubled, increasing from 31% in 1977 to 75.4% in 1997; and (3) The uncontrolled growth of debt among member-states, with the alarming levels reached by Italy, Belgium and Greece (more than 100% of GDP when the euro was launched), undermines the independence and credibility of monetary policy, in that governments may endeavour not to pay high interest rates (or even to avert a financial crisis) by applying pressure on the ECB to monetize their debts with inflationary effects (bail them out). On the other hand, it is argued that the cost of the SGP is low or negligible. Firstly, because the historical experience of the European countries shows that a budget margin of 3% of GDP has always been more than sufficient to correct recessions by means of the so-called 'automatic stabilizers' (wage subsidies, income taxes, etc.).[2] Secondly, because Europe's low growth and high unemployment cannot be remedied by means of a deficit spending policy; rather, they require elimination of market rigidities, especially those of the labour market, and perhaps also reductions of taxation and spending levels kept excessively high by an overly generous social security and welfare system.[3]

Opposition to the SGP by economists has ranged from the radical criticism that it harms the European economy by constraining fiscal policy, to the practical consideration that although a monetary union requires constraints on individual countries' budget policies, for technical reasons these cannot take the form of the SGP.[4] The design of the SGP is misconceived for various reasons. Firstly, because with regard to the crowding-out problem, the deficit/GDP ratio is not an accurate indicator of the absorption of financial resources by a government. Reference should instead be made to the ratio between the deficit and the financial wealth formation of the euro-zone (in which the aggregate incidence of public deficits is relatively slight). This also applies to the bail-out

problem, where the correct indicator is not the deficit/GDP ratio but the dynamic sustainability of the public debt (which ultimately depends on the *future* trends of primary balances, the real interest rate, and rate of growth). Hence an 'excessive public deficit' cannot be determined on the basis of a parameter fixed *a priori* and applied indiscriminately to all countries. Secondly, because resort to sole automatic stabilizers within the 3% margin might be sufficient if the public sector starts from zero deficit, but recessions may persist over time or occur in a correlated manner, so that an output gap may happen when the public sector already faces a 3% deficit and no further stabilization margin is left. Finally, the empirical evidence that the stabilization of the European economies has never in the past required breach of the 3% constraint is anything but well-founded, because it fails to take account of the fact that each individual country's stabilization instruments previously also comprised independent monetary policies and included exchange rate devaluation – measures which are now precluded.[5]

It should be added that following various events in the European and world economy over the past two or three years, the pendulum of the economic debate has swung in favour of the critics of the SGP. For example, no clear correlation has yet emerged between the public deficits of the European countries and the common currency: neither in terms of interest rates, which are still firmly under the ECB's control, nor in exchange rate terms, as demonstrated by the considerable strengthening of the euro (+25% against the dollar) over the past year. Moreover, comparison between the American and European responses to the world-wide crisis since 11 September 2001 lends support to the contention that the worse performance by the European economy has been largely due to the greater rigidity of its monetary and fiscal policies. In the United States, interest rates began to fall in January 2001, and they shed 500 base points in the short-term segment and 250 in the long-term one. In Europe the fall during the same period was only about half as much, which contributed to the euro's appreciation against the dollar.

As regards budgetary policy, as Jean-Paul Fitoussi has recently shown (see Table 10.1), the United States went from a slight surplus in 2001 to a peak deficit of 4.3% of GDP in 2003. The Europe of the SGP started with a modest deficit of 1.7% of GDP in 2001 and then hit 3.8% in 2003 (3.9% for Germany, 4% for France, 2.8% for Italy). If we take the 2001–2004 period to be a recession/recovery cycle, the ratio between average deficit and average GDP during the period reached some 4.7% in the United States, while it remained at 1.2% in the SGP zone. Finally, the 'budget impulse' on GDP has been positive throughout the 2001–04 cycle in the United States but close to nil (or even slightly *negative*) in the SGP zone. To sum up: in the 2001–04 cycle the United States had a mix of more expansionary monetary and fiscal policy, lower *real* interest rates, progressive exchange devaluation, and a more rapid recovery. Although

Table 10.1 GDP related data

Growth of GDP (%)	2001	2002	2003	2004	2001–04
United States	0.3	2.4	2.5	2.9	2.1
SGP Zone	1.6	0.9	0.6	1.3	1.2
Public balance (% of GDP)					
United States	0.4	–2.4	–4.3	–3.2	–4.7
SGP Zone	–1.7	–2.2	–2.8	–2.4	–1.2
Germany	–1.3	–2.8	–3.5	3.9	–3.7
France	–1.3	–1.6	–3.1	–4.0	–4.1
Italy	–1.5	–2.6	–2.6	–2.9	–3.2
Fiscal impulse to GDP (%)					
United States	0.6	2.9	1.8	0.2	1.4
SGP Zone	0.6	0.4	–0.2	–0.2	0.2
Germany	1.5	0.2	–0.6	–0.2	0.2
France	0.3	1.3	–0.1	–0.3	0.3
Italy	0.4	0.3	–0.3	0.1	0.1

Source: Calculations on OFCE (Paris) data from Fitoussi (2003).

this is a partial picture from which hasty conclusions should not be drawn, it nevertheless invites reflection on the technical bases of the SGP.

BETWEEN TECHNOCRACY AND POLICY

Why did the governments that created the euro decide to tie their own hands with the rules of constitutional force laid down by the treaties of monetary union? And, above all why did they do so with 'figures' that have never fully convinced the scientific community and which have proved inaccurate, ineffective, and perhaps damaging when applied in actual practice? These two questions must necessarily be addressed if we are to understand the institutional impasse in which the governments of the Monetary Union now find themselves, and if we are to find possible ways out of it. They are questions that extend beyond the strictly economic domain to merge with political analysis.

As regards the first question – 'why tie your hands with a pact of quasi-constitutional force?' – the historical and doctrinal context should be borne in mind.

The so-called 'Economic Constitution' of the Union (i.e. the parts of the Treaty of Maastricht supplemented by the SGP on the economic institutions) was drawn up in the late 1980s and early 1990s under the strong influence of the 'new political economics' propounded by the Chicago School of Milton Friedman and James Buchanan. This doctrine analyses political and institutional arrangements and their effects on the efficiency of the economic system on the assumption that politicians, heads of government, and senior civil servants are rational actors who take optimal decisions with regard to their *private interests*; but these, of course, do not always coincide with the public interests declared by those actors.[6] This line of thought is imbued with scepticism as to the economic effects of democratic mechanisms. As regards public finance, it maintains that electoral competition and the alternation of governments produce perverse incentives: pre-election fiscal deficits, an inappropriate use of public spending to curb unemployment with inflationary effects, or excessive public debt left to future generations to redeem. Grounded on this mistrust is the idea that the public interest as represented by economic efficiency must be protected against the defects of democratic mechanisms. Essentially two solutions are advocated. First, governments should be exposed as much as possible to the 'constituency of markets', since these are by definition the custodians of economic efficiency and therefore immediately able to detect and punish the inefficient actions of governments. Second, the discretionary powers of governments in regard to economic matters should be restricted and regulated by pre-established rules with constitutional force.

For the sake of brevity, we shall not discuss the problems inherent in the first solution. The second starkly highlights the classic conflict between Technocracy and Democracy, not so much because it sets limits on the exercise of power by freely elected governments (constitutions are created for precisely this purpose and are an integral part of democratic systems) as because the writing of the fixed rules with which governments must abide should be entrusted to a non-political entity, which decides what the economic 'good' is. Indeed, there are those who have talked of the return of the *benevolent dictator*.[7]

It comes as no surprise that academy, the breeding ground of aspiring technocrats, should develop a theory that increases its power. What instead requires explanation is why the highest representatives of the European political class have put this theory into practice with a zeal difficult to find elsewhere in the world.[8] Although it is not easy to find the answer, three important historical contingencies should be pointed out. The first consists, to use James O'Connor's words, in the fiscal crisis of the European model, which erupted in precisely the same years during which the Monetary Union was created. The second has been emphasized by most commentators on the economic treaties, and it concerns the mutual mistrust among the governments which brought the single currency into being. Or better, it concerns the mistrust of the (supposedly) virtuous coun-

tries, principally France and Germany, towards those others suspected of endemic fiscal irresponsibility (for instance, the so-called 'Club Med' or, to use the less exotic but more trenchant acronym, the 'PIGS'– Portugal, Italy, Greece and Spain). The third historical contingency is the reluctance of any government to transfer a significant amount of fiscal sovereignty to a supranational authority.

The answer to the second question – 'why did the governments that created the euro tie their hands in such peculiar manner?' – also lies in the conflict between Technocracy and Democracy attendant on the founding of the Monetary Union. Calculation of the parameters, and their subsequent management at normative level, marked a sort of muddled revenge by Politics. It very soon became clear that, first, the parameters for admission to the Monetary Union, and then those imposed by the SGP, were not only technically unsound for the reasons outlined earlier but were blatantly tailor-made to suit Germany and her option to exclude undesirable partners.[9] This has prompted Paul de Grauwe, one of the leading scholars of the Monetary Union, to comment that

> The imposition of the Maastricht convergence criteria does not seem to have been based on economic analysis [… They] serve the purpose of restricting membership in the monetary club and keeping it small (1995, pp. 484, 487).

Subsequently, however, when the day of reckoning came for admission to the Monetary Union, the political stand-off over who should be admitted and how was resolved by resorting to the original scheme set out in the Delors Report[10]: a 'broad' monetary union which would serve as the basis for political union (not vice versa). Countries ineligible on technical grounds – most notably Italy and Belgium because of their excessive debt/GDP ratios – were granted admission by inventing *ad hoc* graduality clauses. Thereafter it was clear – or should have been to the most acute observers – that the SGP was no more than a Maginot Line: an ambitious but technically flawed structure, easily overwhelmed by the Politics against which it should have been a bastion. Nevertheless, it remained in place as a major hindrance against smooth economic policies in the field of monetary union. Hence the Ecofin's political act of November 2004 came as no surprise; rather, it was the full-blown manifestation of the contradictions inherent in the SGP.

Few still believe that the SGP has a future in its present form; but there are many who are willing to credit it with making the Monetary Union possible and for forcing fiscally irresponsible politicians and voters to undertake painful, yet necessary policies to restore healthy public finances and ensure their better management. We have no objections on this score, but we submit that it was obtained with significant damage to the already fragile structure of the Community's institutions. The severe institutional crisis, as well as of legitimation

and public popularity, now besetting the Monetary Union requires critical rethinking, and above all a commitment not to go on using Europe as an excuse for 'reform-mongering'.[11]

WHAT IS TO BE DONE?

The way out the impasse should be neither in the direction of Ecofin nor in that of the Commission. On the one hand, the Commission is undermining its credibility being compelled to defend a rule enshrined in the treaties, but no longer defendable in either form or substance. The minor adjustments to the 3% rule of March 2005 (the break-up of certain budget items, the so-called 'golden rule', escape clauses, etc.) are not acceptable solutions, both because they are difficult to implement and, especially, because these adjustments do not remove the intrinsic flaw in the SGP: that it is a perverse and fragile blend of technical rules with political purposes (and political choices) concealed behind a façade of 'purely' technical rules. On the other hand, Ecofin seems prey to a confused desire by Eurosceptic governments to assert the Politics of individual states. What these governments want is a Europe *à la carte*, something little more than a broad free trade area in which a common currency circulates subject to a minimum of economic coordination, and in which nothing has legitimacy beyond the will and power of chancelleries; a Europe, as Commissioner Monti worriedly put it, which is pre-war if not nineteenth-century.

The future of the Monetary Union instead requires that Politics take the responsibility for a profound reform of the Pact among the constituent governments, in order to rule on budget prerogatives and institutional relationships with the monetary authority within a framework of appropriate technical instruments. The best solution, we believe, would be resumption of the scheme set out by the Delors Report: the creation of a European Confederation Treasury with increasingly broad prerogatives and sources of legitimacy, autonomous decision-making capacity, and its own budget, financed partly by the member-states and partly by the issue of long-term bonds.[12] A proper European Confederation Treasury would pursue Community-wide *stabilization* and *structural policies* with high 'externalities' (defence, large infrastructures, intra-Community income redistribution, etc.), leaving individual governments with lighter budgets and tasks focused on *local policies*. This would achieve the functional distribution of fiscal responsibilities that one observes in the most advanced federal systems. We may call this the 'long-period federalist solution', since it entails a radical political shift towards a supranational transfer of powers which many states and many political parties are not yet willing to contemplate.[13] In the shorter term, however, reform of the rules concerning national budgetary policies is at any rate necessary, and we maintain that it can be un-

dertaken without any change to the distribution of powers. Its form can be summarized as follows, where the focus is naturally more on principles than on details.

1. *A Budgetary Pact*. Given an entirely independent central monetary authority, the less fiscal and spending powers are transferred from the periphery to the centre the greater the need for a 'Budgetary Pact' (BP). The principles underlying this Pact should be the following: (a) Over the short-medium period, the aim should be compatibility of national budgetary policies with European-level macroeconomic targets (regulation of aggregate domestic demand, inflation rate, euro exchange rate); and (b) In the long period, the aim should be financial sustainability and convergence among the dynamics of national public debts.[14]

2. *The BP should be intergovernmental and with constitutional force*. It is evident that the BP should reflect the political will of the legitimate custodians of fiscal and spending powers – national governments. On the other hand, steps should be taken to prevent reciprocal trust among the member-states from being undermined by excessive scope for discretion, manipulation or collusion in interpretation and application of the Pact. It is legitimate to expect that once the Pact has been signed it should apply to both large and small countries regardless of changing political attitudes and independently of changes in the alliances among countries with their respective zones of influence. This requires that the Pact, in its *form*, should be superordinate to the powers that national governments have in Ecofin.

3. *The BP requires a European Budgetary Authority*. A constitutional pact is nothing but a contract. A contract, however, is bound to be incomplete unless an omniscient actor (a Technocrat?) is able to foresee, describe and prescribe the 'right' behaviour for all the parties concerned in all possible future circumstances. The failure of the SGP demonstrates that this doctrinal approach is fallacious in an uncertain world. The enforcement of a pact of this kind neither requires nor justifies the imposition of quantitative and fixed budget parameters equal for all countries and in all circumstances. Instead, as happens in all areas regulated by constitutional rules, it requires an authority for the assessment and authorization of the behaviour admissible in given circumstances.

This approach does not entail a revolution in the Monetary Union. As regards the monetary authority, it already exists *in nuce*. The Treaty of Maastricht, in fact, lays down the general principle that the primary objective of the ECB must be price stability (article 105), while the ECB itself has full discretion in *defining, measuring and quantifying* this objective in relation to specific technical and economic circumstances. Likewise in regard to budgetary policy, the Treaty establishes the general principle of financial stability and the obligation to avoid 'excessive public deficits'. This is already a norm with constitutional force, and it is only necessary to give symmetry on the monetary and budgetary sides by

replacing inappropriate predefined quantitative constraints with a Budgetary Authority which (a) specifies the general principles of good practice with which governments must comply (for example, in accordance with the objectives set out at point 1 above), and which should engender appropriate behaviour once the specific circumstances are known; and (b) is empowered to evaluate and authorize budget programmes admissible in those circumstances.[15]

4. *Legitimacy and powers of the Budgetary Authority*. The Budgetary Authority must have political legitimacy but it must also be independent from national governments (Ecofin). It should be a branch of the Commission, and it should be under the responsibility of the President of the Commission and of a special commissioner (in the Giscard-Amato draft, the President of the Commission is elected by the European Parliament and therefore has political legitimacy). The objections raised against this solution by the defenders of the *status quo* are largely of practical nature. Without rules that are simple, clear and equal for everybody, they claim, the Authority would be faced with an impossible task.[16] These objections are unconvincing, however. Firstly, the idea that it would be simple, clear and fair to apply the 3% rule has proved illusory. Secondly, once the status and mission of the Budgetary Authority have been defined (see point 3), its procedures and competences would not greatly differ from what now happens under the SGP, with governments obliged to submit detailed multi-year budget plans scrutinized after an enormous amount of preparatory work and negotiations among national ministers, Ecofin and the Commission. Thirdly, once the general criteria (for example those outlined at point 1) have been defined, establishing whether or not a national budget entails an 'excessive deficit' is not an insurmountable problem, given the technical instruments and expertise available to the Commission, which are certainly not inferior to the Monetary Fund or the rating agencies that normally produce evaluations of this kind. Finally, there is no basis to the argument that it would be politically untenable to allow (or not to allow) different countries to have different budgets. On the contrary, it is well known that the crucial aspect of a monetary union is deciding how countries should be allowed to deal with asymmetric shocks, and that this requires diversified budgetary policies, and that the less correct national budgetary responses are, the greater the damage for everyone. Individual governments, if they are rational and have their own national interests in mind, should concern themselves with the correctness and uniformity of the rating procedures, not with the size of their partners' budgets year by year.

To conclude, success of the Monetary Union will eventually require a fiscal and political union, thus resuming the route marked out by the Delors Report. Immediate creation of a Federal Treasury is, however, not necessary to produce a better mix of national budgetary policies with single monetary policy. We believe it feasible to create a Budgetary Authority within the Commission whose task is evaluating and approving the multi-year budget plans of member-states,

according to the general principles set out in the Treaty of Maastricht. The powers of the Budgetary Authority can be defined with variable geometry and with increasing range so that they are attuned to the orientations and federalist maturation of national governments. Notwithstanding the numerous difficulties to overcome, and the time-frame of Politics, on this occasion the first steps must be made in the right direction.

NOTES

1. Europe split not only among institutions (Council and Commission) but also among countries. Denmark, Austria, Finland and Spain sided with the Commission, while votes in favour of Germany and France were cast not only by Italy (without whose vote the qualified majority for the document would not have been reached) but also the United Kingdom, which does not belong to the monetary union, and a number of small countries like Ireland and Protugal, which for some time have shown impatience with the Commission's rigour.
2. A frequently cited study shows that, on average in the European countries, a 1% fall in GDP produces an 'automatic' deficit of 0.5%. From this one deduces that, starting from a balanced budget, the 3% constraint would be reached in the presence of a 6% recession, an event considered to be highly unlikely. Cf. V. Tanzi (2003).
3. For detailed treatment of these topics see M. Buti and A. Sapir (1998).
4. The cover of *The Economist* of 16 December 1996 carried the title *The Wrong Design* with reference to the GDP.
5. For detailed discussion of these topics see e.g. F. Farina and R. Tamborini (2002).
6. For this interpretation of the treaties, see e.g. T. Padoa Schioppa (1994). An introduction to new political economics is provided, amongst others, by T. Persson and G. Tabellini (1990).
7. Cf. Jean-Paul Fitoussi (2002).
8. The European economic treaties restrict the discretionary powers of governments to a much greater extent than is the case in the United States. It is true that some of those states are subject to the balanced budget obligation or to indebtedness restrictions, but this happens in the context of a strong federal fiscal system, which assumes the functions of both stabilization and redistribution.
9. Among the first to point this out were W. Buiter, G. Corsetti and N. Roubini (1993). In short, the stability condition of the debt/GDP ratio states that the deficit/GDP ratio should, by approximation, equate the target debt/GDP ratio times the nominal GDP growth rate. The hisorical nominal growth rate in Germany was about 5% and the debt was 60% of GDP, which would require a deficit/GDP ratio of 3%.
10. This Report, issued in 1989 by the Commission when Delors was its President, outlined the route that would lead to the Treaty of Maastricht and creation of the euro.
11. The term was coined by A. O. Hirschmann with reference to the Latin America of the 1960s.
12. Today, the EU budget: (a) by law cannot be more than 1.25% of the EU member GNPs; (b) *de facto* never exceeds 1%; (c) by law cannot be in deficit.
13. However, this is not to say that a number of common principles on taxation are not already under discussion, like the harmonization of tax bases, especially for company taxes.
14. An example is the proposal set out in F. Targetti (2003).
15. For the sake of completeness, it should be pointed out that various commentators have correctly complained about the absence of an institutional body to which the Central Bank is accountable for *how* it applies the price stability principle. A persuasive suggestion is that the monetary and fiscal authorities should act as checks and balances for each other's policies. Moreover, a number of recent studies have shown that agreement between the monetary and the fiscal authority on the desired levels of inflation and GDP would leave each of them with

full discretion in implementing their respective stabilization policies, with globally optimal results; see A. Dixit and L. Lambertini (2001), R. Tamborini (2003).

16. Cf. M. Buti et al. (2003).

REFERENCES

Buiter, W., Corsetti, G. and Roubini, N. (1993), 'Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht', *Economic Policy*, **16**.

Buti, M., Eijffinger, S. and Franco, D. (2003), 'Revisiting the Stability and Growth Pact: Grand Design or Internal Adjustment?', *European Economy*, Economic Papers, no. 180.

Buti, M. and Sapir, A. (1998), *Economic Policy in EMU*, Oxford: Clarendon Press.

De Grauwe, P. (1995), 'Alternative Strategies Towards Monetary Union', in Papers and Proceedings of the European Economic Association, *European Economic Review*, **XXXIX**.

Dixit, A. and Lambertini, L. (2001), 'Monetary-Fiscal Interaction and Commitment versus Discretion in a Monetary Union', in Papers and Proceedings of the European Economic Association, *European Economic Review*, **XLV**.

Farina, F. and Tamborini, R. (2002) 'Le politiche macroeconomiche di stabilizzazione in Europa nel nuovo regime di unione monetaria', in F. Farina and R. Tamborini (eds), *Da nazioni a regioni. Mutamenti strutturali e istituzionali dopo l'ingresso nella Unione Monetaria Europea*, Bologna: Il Mulino.

Fitoussi, J.P. (2002), *La régle et la choix*, Paris: Editions de Seuil.

Fitoussi, J.P. (2003), 'Lo scenario internazionale e il governo economico dell'Europe', lecture given at the Faculty of Economics, University of Trento, 10 November 2003.

Padoa, Schioppa T. (1994), 'Governo dell'economia e contesto europeo', *Il Mulino/Europa*, **1**(2).

Persson, T. and Tabellini, G. (1990), *Macroeconomic Policy, Credibility and Politics*, Chur: Harwood Academic Publishers.

Tamborini, R. (2003), 'The 'Brussels consensus' on macroeconomic stabilization policies in the EMU. A critical assessment', Paper presented at the First EUI Alumni Conference on *Governing EMU: Political, Economic, Historical and Legal Aspects*, European University Institute, Florence, 3–4 October.

Tanzi, V. (2003), 'Role and future of the Stability and Growth Pact', Paper presented at the 21st annual conference on *The future of the Euro*, Cato Institute, Washington, DC, 20 November.

Targetti F. (2003), *Le conseguenze economiche del governo Berlusconi*, vol. 6, 'Patto di stabilità e bilancio europeo', Milano: Nuova Iniziativa Editoriale.

11. Narrowing the options: the macroeconomic and financial framework for EU enlargement

Heather Gibson and Euclid Tsakalotos

INTRODUCTION

Keynesian economists, in the tradition of Nicholas Kaldor, Paul Davidson, and Tony Thirlwall, have very strong opinions about the economy and economic policy.[1] But more importantly for the purposes of this chapter, it is also the case that they have sought to promote a framework for economic decision-making that allows for the clash of ideas to have some impact on outcomes. That is to say, they have not seen democratic accountability and discretion, as is the case with many of their 'classical' and 'new classical' opponents, as a problem to be overcome through various devices such as independent central banks, budgetary rules and so on. Rather they have seen democratic processes as part of the answer to good economic policy-making. In part, this stems from their view on the inherent instability of the market economy, requiring as it does appropriate remedies. Moreover these remedies are unlikely to be the same for all economies for all time. But, in part it also derives from an optimistic stance about the human condition – it should not be beyond people to devise a better institutional framework to deal with economic instability, promoting economic growth as well as addressing issues such as poverty and income inequality.

Over the last twenty years, this admirable stance has been on the defensive. Nowhere is this clearer than in the manner in which European integration has been promoted. At the level of official EU discourse, with all the emphasis on subsidiarity, and more recently on benchmarking and peer group pressure to find best practices, there is supposed to be no commitment to a particular model of capitalism that would guide new members. At the level of academic discourse, and more generally political ideology as well, it may appear that the high tide of neoliberalism has been reached, implying that the new members (the EU10), not to speak of the older members (the EU15), are quite free to choose both the economic policies appropriate for their economies and also where exactly they would like to develop on the axis between the two poles of a liberal capitalist

economy and a more institutional economy (in which the market and hierarchical firms are more integrated with other mechanisms of economic governance such as states, associations and communities).[2]

The argument of this chapter is that reality is somewhat different. The institutional framework and the policies of the EU are far from neutral with respect to the two poles. The pressure on new entrants to keep as close as possible to the liberal pole, for both policies and institutions, will almost be impossible to resist. To make this argument we start by examining the macroeconomic framework and then go onto consider the financial framework.

THE MACROECONOMIC FRAMEWORK

The framework of macroeconomic policy for acceding countries is fairly restrictive, and individual countries are unlikely to have much room to decide between policy options on the basis of what is best for their own economies. At the very least, this implies that the EU10 will probably have a 'bumpy' ride on the way to EMU (and perhaps even after euro area entry). We can illustrate this by looking at three areas of policy: first, exchange rate and monetary policy; second, fiscal policy; and, finally, the so-called Broad Economic Policy Guidelines (BEPGs), which cover not only monetary and fiscal policy issues but also employment and other policies aimed at structural change.

Exchange Rate and Monetary Policy

Entry into the EU entails the expectation that at some point in the future the acceding countries will join the euro area. One of the conditions laid down in the Maastricht Treaty for euro area entry is ERM membership for a period of at least two years. ERM II, which succeeded the old ERM on 1 January 1999, is highly asymmetric. It is a 'hub and spokes' system with the euro area as the 'hub' to which the acceding countries, as the 'spokes', will tie their exchange rate. The ±15 per cent fluctuation bands give some flexibility and are designed to accommodate countries at very different stages of convergence. Realignments are permitted and are joint decisions taken by the ECB, other central banks in the system and finance ministers.[3]

Whilst the system is more flexible than the old ERM (at least in its pre-1993 version), and this is useful since it allows those in the EU10 which are close to euro area convergence to operate a tighter exchange rate target while those that are further away have more time to get used to operating within the rules of the game, there are two main reasons why it does not help to reduce the bumpiness of the EMU path. First, ERM II is asymmetric by design, with the onus of adjustment completely on the EU10 countries themselves. This has arisen in the

attempt to build and preserve ECB credibility by giving it complete independence to pursue its policy of price stability. Although the ECB is required to intervene at the margins, it has the right to limit (or even suspend) intervention if it judges that the currency under attack is misaligned. Thus countries can expect little help to support their currencies through difficult times if the ECB fears such support may undermine its ultimate goal – price stability in the euro area. Moreover, disturbances in foreign exchange markets may well be magnified by such a stance, since it is well established that coordinated intervention is more successful than uncoordinated intervention.[4]

The second reason why the system does not provide much support for the EU10 countries is the lack of genuine policy cooperation. ERM II entails what the agreement terms 'policy co-ordination'. This effectively can be translated as the EU10 countries coordinating their policies with the given policies of the euro area. Cooperation, by contrast, would involve discussions leading to compromises where both sides frequently move from their initial positions. One example of cooperation, which could be very useful in supporting the EU10 countries' efforts to join the euro area, is a reinforcing interest rate change to support a currency that is under attack. Yet such policy cooperation has been strongly resisted by the ECB in all its policy statements on ERM II. Again, the aim is to preserve the independence of the ECB and avoid situations where it might have to subordinate its goal of price stability to some other policy objective.

What are the implications of this for the EU10 countries that wish to become euro area members, particularly in light of the fact that capital controls are to be removed on entry into the EU? In earlier work (Gibson and Tsakalotos, 2004) we examined the capital flow experience of the EU10 countries along with Bulgaria and Romania (two of the accession countries) with a view to shedding some light on this question. This entailed the estimation of two models: the first to examine the determinants of the level of capital flows and the second to focus explicitly on periods of downward speculative pressure. From this exercise we drew a number of conclusions. The evidence suggests that capital inflows and outflows, including episodes of speculative crisis, have been of a significant magnitude in these countries, at least as great as those experienced by the southern European economies. The comparison is interesting since it is well documented that both large capital inflows and outflows complicated the conduct of monetary policy in Greece, Portugal and Spain in the run up to euro area membership (Gibson and Tsakalotos, 2002).

On the issue of the factors underlying these flows, we found that the impact of domestic factors, such as inflation and growth, is rather weak. While the adoption of a currency board or a pegged regime tends to increase capital inflows, what is of far more importance are external factors, notably in our period the strong effect from the Russian crisis in 1998. This was also the case when

we looked at the probability of a speculative attack during periods when these countries pegged their exchange rates in the 1990s. The results again confirmed the importance of contagion.

The implications of these findings are clear. Exchange rate targeting within ERM II may well be undermined by capital flows. Indeed, it is not enough for a country to ensure that its domestic policies are 'sound'; the significant effects of contagion suggest that even with sound domestic policies, ERM II membership could well prove to be a bumpy ride. Yet this ride could be made smoother by a greater commitment on the part of the ECB to intervene to support central rates and/or by more genuine policy cooperation between the euro area and the EU10 countries. Alternatively greater autonomy for macroeconomic policies could be provided if emerging economies were allowed to employ capital controls. The international financial crisis in the late 1990s, the fact that influential economists such as Krugman and Stiglitz were arguing in favour of reform,[5] and the election of centre-left governments in a number of influential European countries may have constituted a window of opportunity to change the rules of the game within the international financial system. But this moment has now passed and it seems unlikely to return in the current intellectual economic climate within EU policy circles.

Fiscal Policy

If monetary and exchange rate policies are likely to be constrained by the framework of an asymmetric ERM II, then fiscal policy is likewise limited by the Maastricht Treaty and later, within the euro area, by the Growth and Stability Pact. The rules laid down in the Maastricht Treaty regarding budget deficits and the level of general government debt reflect not only the desires of the mostly centre right governments in the 1980s and early 1990s, but are also supported by the prevailing consensus among economists. The recipe of the new classical and credibility literature[6] lies in tying the hands of government, not only in the area of monetary policy, but also with respect to fiscal policy. The aim is to ensure credibility; the underlying assumption being that we have more to fear from destabilising government action than from the instability endogenous to the market economy itself. Moreover within this perspective, fiscal consolidation is seen as being compatible with high employment and growth, through what might be called 'expansionary fiscal consolidation'. A reduction in budgetary deficits, allowing for a more balanced mix between monetary and fiscal policies, could lead to lower interest rates, higher investment and greater aggregate demand.[7]

Of course, the conditions for this operating are rather strict,[8] and it is far more likely that such consolidation leads, along traditional Keynesian lines, to pessimistic expectations about the future, thereby worsening the recession. The

mainstream account seems to suggest that once consumers and businesses see the tight fiscal environment they will become more confident about future economic conditions. What most worries economic agents, in this account, is government spending in projects of uncertain worth leading to more bureaucracy and waste, higher interest rates and future increases in taxes. Since demand plays little role here, it is assumed that if the private sector is not investing in projects, it must be the case that there are no profitable projects in which to invest – unless of course there are supply side reforms (of which much more shortly). This is a more or less direct replay of the arguments about 'psychological crowding out' that were made during the inter-war recession.[9] But now as then, it does not seem that the adoption of the principles of 'sound finance' has led to any obvious buoyancy in the psychology of Europeans. Indeed it is far easier to argue that the Stability Pact as a device that was promoted in order to promote credibility and stability has itself led to increasing pessimism. Restrictive fiscal policies significantly dampened economic activity in the run up to EMU (see Fitoussi, 1997 and Modigliani *et al*, 1998). Nor can it be argued that this was just a transitional phenomenon. As De Grauwe (2004) has recently argued, 'With each economic downturn, budget deficits start increasing[10] and at that point the Pact causes Brussels to issue stern warnings about unsustainable budgets, irresponsible governments and other nasty things. Thus the pact creates a psychology of crisis with each downturn in activity, thereby undermining confidence in the recovery. In a perverse way, the pact has become a pro-cyclical force intensifying the recession'.

Moreover what the obsession with inflation, which underlies the Pact, ignores is the fact that some inflation is integrally tied to the growth process, that it can stimulate investment by reducing real interest rates and raising prospective yields on investment, and that it is useful when there is real wage flexibility but nominal wage inflexibility (see Thirlwall, 2001). However whenever European policy makers have been faced with the choice of provoking unemployment or disappointing the rentier, they have rarely been paralysed by indecision.[11]

Nor is there any role in the above conception for taking account of Europe's important relationship with world demand. Forder (1999) asks us to consider the following realistic scenario. Say that South-East Asia is in an economic downturn and needs to increase its exports to the rest of the world. Assume also that the US, because of a period of high growth is unable to expand further. This leaves Europe, which can either import more by exporting less, with deflationary consequences, or it can coordinate a European-wide fiscal expansion. Without such an expansion the economic downturn in the East would soon lead to worsening economic conditions in Europe. Thus this European expansion would seem, given the recent record of sluggish growth and high unemployment, to be both desirable and possible. But the EU has not developed the means to respond to such an eventuality. Of course such considerations hardly exhaust the

potential for the EU to intervene in the direction of a more expansionary international economic environment. The fact that the EU cannot afford to ignore world demand suggests that what would be of more lasting value is a reform of the international financial system along the lines promoted by Keynes in 1943, and his followers subsequently (see Thirlwall, 1999).

Given all the above it is not surprising that for the EU10 countries meeting the fiscal criteria is one of the biggest challenges they face.[12] The claim that an upper limit for deficits of 3 per cent gives plenty of room for automatic stabilisers to work is looking less secure in the light of developments in the recent downturn. For transition economies whose economic cycles tend to be of greater amplitude, this will almost certainly prove to be constraining. Furthermore, the EU10 countries, which are also completing the transition to a market economy, have a need for investment in the areas of infrastructure and education and training, which at the very least require financial support from the government. Yet government capital expenditure is not treated under the Maastricht Treaty in any way differently from current government expenditure that is not expected to yield a return in the future. Thus on fiscal policy too, the options available to the EU10 countries are limited.

BROAD ECONOMIC POLICY GUIDELINES

Finally, there is the umbrella framework of the BEPGs, which are discussed in the macroeconomic dialogue that seeks to ensure the appropriate mix between monetary and fiscal policy. The BEPGs aim at coordinating policy between all EU member states (be they in the euro area or not) and they include both directions for macroeconomic policy and structural policies. In terms of the macroeconomic framework, the emphasis is on monetary policy being committed to price stability within the context of an independent ECB and fiscal policy as envisaged in the Maastricht Treaty and the Stability and Growth Pact. Additionally, emphasis is placed on wage developments that are expected to contribute to macroeconomic stability. In particular, to quote from the BEPGs agreed in 2003, the social partners should '[e]nsure that nominal wage increases are consistent with price stability and productivity gains. In particular, wage developments should remain moderate in the context of a possible cyclical recovery in productivity or oil-price-hike-induced increases in inflation to allow for a restoration of profit margins so as to underpin job-creating investment growth' (European Commission, 2003, p. 5). A more full-hearted return to the theories and priorities of the classical economists, that Keynes took as his starting point for his critique in the *General Theory* (see Thirlwall, 1999), would be difficult to imagine. As long as there is low inflation and sound finance, labour flexibility can be relied upon to ensure macroeconomic bliss.

It is important to see how restrictive this overall approach is. In the first place there is very little democratic debate about central aspects of economic policy-making. In much of the credibility literature it is often assumed that no issue of democratic accountability is as at stake since monetary policy does not affect real variables in the long run. Issues of a technocratic nature, it is implied, can be left to technocrats. Such a conception leaves much to be desired. Even if we accept, for the sake of the argument, that there is money neutrality in the long run, there may still be important issues for democratic deliberation on paths back to equilibrium (often entailing different distributional trajectories), on the effects of interest rates on external balance, and much else besides (see Forder, 2000).[13] Within a broader framework, it is possible that people may want to trade-off some price stability for greater growth and less unemployment, but they are not given that choice (Thirlwall, 2000). It is truly a sign of the times, of what Paul Krugman has termed 'diminished expectations', that so many economists see so little role for democratic procedures with respect to economic issues.

In this sense, the increasing emphasis given to macroeconomic dialogue in EU policy documents (see Tsakalotos, 2001) is misleading and does not signal a change in direction of macroeconomic policy. Rather, it is used for the EU bodies to communicate effectively their views on the course of the macroeconomy and hence what they believe might be the appropriate response of the social partners in their wage bargaining process.[14] It would be wrong to consider the role of the social partners in this conception as a kind of corporatism at the European level. While there have been important recent national developments in countries such as Portugal, Belgium, Ireland and Spain, the content of this bargaining is much narrower than in traditional corporatist arrangements and the focus is on 'credibility bargaining' to produce a 'national institutional lock-in' (Teague, 1999, p. 39). After EMU, the focus is on ensuring that the various economies do not lose their competitiveness,[15] becoming backwaters within the euro area. All this is far removed from the broader agreements and 'political exchange' of the social corporatist model. If the legacy of the macroeconomic straightjacket as a response to inflationary pressures can be seen as a legacy of the dominant right-wing governments of the 1980s and early 1990s, then centre-left politicians have been unwilling to take up more corporatist solutions to the problem of inflation in the subsequent period.[16]

Rather both centre-left and centre-right governments, at home and at the European level, have sought to underline their continued commitment to low inflation with a more determined effort to improve employment performance through microeconomic policies on the supply side.[17] While 'Keynesian theory teaches that wage flexibility may not be an efficient adjustment mechanism in the case of demand shocks' (Thirlwall, 2000), unemployment is now, as in the inter-war period, seen by European policy makers largely in terms of the inflex-

ibility of the workers themselves. Thus macroeconomic demand management plays no role at all within the European Employment Strategy, with all the emphasis being on increasing flexibility and entrepreneurship. To be sure, market failures are indicated, and policies promoted to rectify them not least in the area of education and training, but the amount of discretion allowed to the social partners is very limited.

In the post-war era, social democratic Keynesianism and/or corporatist arrangements can be seen as providing an institutional basis for the accommodation between labour and capital.[18] On the other hand, the doctrine of 'sound finance' can only be seen as an *ex parte* intervention. Without a full-employment guarantee, labour faces the stark choice of 'pricing itself back into the market' (a strategy which may not even work) or unemployment, something which clearly weakens its bargaining power. Moreover, the macroeconomic framework of 'sound finance' confirms the power of financial markets to give a verdict on the appropriateness of economic policy and, given the extent to which such market actors represent the most liberal type of capitalism, these actors are hardly likely to look favourably on more institutional forms of capitalism (Crouch, 1997, p. 358). And it is to the financial markets that we now turn.

THE FINANCIAL FRAMEWORK[19]

Financial liberalisation has been on the agenda since 1945, but a firmer commitment to more liberal financial markets was associated with the neoliberal turn in economic policy-making in the 1980s. It is true that the term financial liberalisation encompasses a number of, often very distinct, issues. Indeed in a previous paper (Gibson and Tsakalotos, 1994) we argued that the economic literature concerning financial repression rather tended to ignore important distinctions. Thus the practice of holding down interest rates, with consequences for savings and growth, is a very different issue from the existence of allocated credit for particular industries, the potential role for a state development bank, or a less market based and more bank based system for corporate control. By lumping all these issues together as aspects of financial repression, supporters of financial liberalisation were not required to give detailed arguments for their support for financial liberalisation across the board – there was, in other words, a real danger of throwing out the baby with the bathwater.

EU policy in the financial sector only makes sense on the assumption that liberalised financial markets will ensure that sufficient funds will go were they are most needed. The EU10 countries have not been offered a choice of policies and institutions most relevant to their own needs, but must make do with the rather narrower menu offered in a more liberalised environment. It is as if countries such as Poland and Bulgaria have nothing to learn from the recent

development trajectory, in a much less liberalised environment, of a country such as South Korea, let alone the earlier development of countries such as France and Germany. As we shall see the prospects for convergence with the rest of Europe in this context are not particularly promising. The fear is that we have another case of what Chang (2002) has recently described as kicking the ladder away.

EU policy in the area of financial services clearly places emphasis on the promotion of markets rather than the creation of a more institutional economy. The Financial Services Action Plan, published in May 1999, has three main objectives: first, to create a single wholesale market; second, to ensure open and secure retail markets; third, to develop state-of-the-art prudential rules and supervision. The first of these objectives is that which is of most interest here, since it involves a series of policy measures designed to allow firms to raise finance at low cost from EU-wide markets, to provide investors and intermediaries with a single point of entry to all EU markets, and to permit investment services to be supplied on a cross-border basis. A related set of policies aims specifically at new, high-growth firms is contained in the Risk Capital Action Plan (RCAP).[20]

The policies contained in the FSAP and the RCAP relate to debt securities, publicly issued equity and venture capital. The rationale is the promotion of financial *markets* as providers of finance capital to firms rather than financial *intermediaries* such as banks. Some of the policies contained in these plans relate to question of the conduct of corporate governance. Although the Commission has tried to duck the issue of market versus bank based systems, it has arisen in respect to the FSAP. Thus, for example, the Directive on Takeover bids represents an attempt to introduce more market-oriented methods of corporate governance by facilitating hostile takeover bids.[21] These developments are seen as crucial in promoting economic growth and convergence and hence are of particular interest to the acceding countries.

That financial markets are becoming more important is clear from the rapid growth of securities issues since the creation of the euro area. Gross issues of short-term and long-term securities denominated in euros by non-financial corporations in the euro show a sharp rise since the beginning of 1999 (see Gibson and Tsakalotos, 2003). Similar rises are evident for monetary and non-monetary financial corporations; indeed in the former case the overall level of issues is much higher. Additionally there has been a rise in the importance of the euro as a currency of issue. Euro issues have been approaching those of dollar issues in various markets such as those for Commercial Paper, Floating Rate Notes and fixed interest rate bonds. An interesting point to note is that these trends have not been affected by recent downturns in global economic activity.

Since there seem to be advantages and disadvantages to raising funds both from markets and from banks,[22] why not simply allow both to develop? In other

words the EU, by seeming to allow both to develop, could be seen to be acting neutrally between liberal and institutional capitalist regimes. We would argue that this is a particularly misleading view of the way in which economic systems in general, and financial systems in particular, work. Economic systems are not like supermarkets where different products can simply be picked off the shelf and placed in one's trolley. Rather different types of systems have their own logic. More market-based systems come with a particular culture of production and certain social relations which tend to be short-term and are characterised by an immediate satisfaction of wants. In the case of financial systems, there are a number of ways in which the promotion of markets will sit uneasily with a desire to promote the kind of system observed in Germany or Japan, characterised by longer-term relationships.

First, even though internal finance is so dominant, markets do tend to create a situation that Grahl (2001) has described as 'the tail of external finance wagging the dog of internal finance'. The existence of market-based financing plays a larger role in defining the opportunity costs for borrowers and lenders. Borrowers in industry, for example, take the external cost of capital as a hurdle rate for investment projects; lenders to industry look at other assets to provide some kind of benchmark return against which they judge a particular project. Such a role for markets could be beneficial if financial markets were operating to ensure that resources were allocated to the most productive uses. But, even if we accept that stock markets exhibit information arbitrage efficiency, it can still be doubted that they exhibit what Tobin (1984) has called 'fundamental valuation efficiency', that is, that prices in financial markets always reflect the future income stream obtained from holding that asset. Thus, Shiller (1981) has argued that share prices are far too volatile to be justified by movements in fundamentals. There is a large literature (Camerer, 1989), building on Keynes' characterisation of the stock market as a beauty contest, which discusses the potential for markets to generate bubbles, either through rational or irrational behaviour. Thus the promotion of markets may well act to the detriment of the performance of the real economy.

Second, financial markets tend to create a 'culture of deal making' (Tobin, 1984), which drives out a culture based on implicit contracts and long-term commitment. Tobin has argued that 'we are throwing more and more of our resources, including the cream of our youth, into financial services remote from the production of goods and services, into activities that generate high private rewards disproportionate to their social productivity' (p. 294). Too often such activities simply create private and not social benefits, and are associated with the redistribution of wealth rather than its creation (Stiglitz, 1989). But the existence of such private benefits in excess of social benefits creates an incentive for overinvestment in working out the true value of companies.

Third, there are a number of theoretical arguments that suggest that market-based finance may well drive out bank-based finance. Thus, it might be more

difficult when well-developed markets exist to promote long-term commitment financing which can help to reduce the liquidity constraints which firms face (Petersen and Rajan, 1994) and can help firms through bad periods (Mayer, 1994). Mayer (1994) argues that commitment lending is very vulnerable to competition from market-based sources of finance. If committed lenders are to see firms through both good and bad times, then they expect the lower returns received in bad times to be made up for by a better return in good times. However, in good times it is easier for firms to get finance from markets, and since this will be cheaper the incentive for firms to renege on the commitment relationship is great.

A final theoretical argument that suggests that markets can destroy more optimal cooperative arrangements relates to the operation of the market for corporate control. Shleifer and Summers (1988) argue that takeovers tend to result in implicit contracts and cooperative agreements between the various stakeholders in a firm being broken because they lead to a substitution of behaviour based on an immediate benefit for cooperative long-term trust relationships which are beyond the immediate equivalence of benefits. If one agrees with Kay (1993) that firms' relationships with their customers, suppliers, workers and financiers are critical to the 'architecture of corporate success', then a greater role for the market, which destroys such relationships, is detrimental to the performance of the real economy.

Thus the 'success'[23] of American-style financial institutions is not *ipso facto* a sign of greater efficiency. Such financial institutions have benefited from the global financial liberalisation of recent years that has been a process where political aspects have surely been as important as considerations of economic performance. In this context financial systems based on the German-Japanese model have come under pressure, as Grahl (2001) has recently argued, but this is because of the nature of the external environment, of which the types of financial polices being promoted by the EU stand out as paradigmatic cases.

How well the new financial system will serve the EU10 economies as well as other possible new EU members remains to be seen. But there are clear grounds to be sceptical.[24] Thirlwall (2000) himself has clearly outlined some of these with respect to the likely effect of further European integration on regional divergences within Europe. His pessimism is based on a set of overlapping concerns that have deeply influenced his work: the operation of cumulative causation as opposed to the neoclassical faith in financial flows naturally smoothing out regional differences; the existence of a balance of payments constraint on growth; the role of non-price factors in competitiveness; and the importance of manufacturing industry to growth and development. Crucial to this worldview is the idea that the market can indicate problems but cannot automatically provide the appropriate responses. While not neglecting the role of structural funds, he argues that the programmes, which are presently severely

underfunded, 'need to create sustainable comparative advantages in disadvantaged areas, by developing clusters of inter-linked activities which can generate self-sustaining growth and weather the vicissitudes of the trade cycle […] In my view what regions and countries require as far as possible are individually tailored economic and social policies to suit there own particular needs' (*op. cit.*). It is precisely what Conservative economic policy, with its faith in markets, failed to achieve in the 80s – by neglecting manufacturing, and the role of non-price factors, this strategy failed to raise 'the achievable growth rate' (McCombie and Thirlwall, 1992).[25]

The argument of this chapter is that the financial and macroeconomic framework of the EU at present makes it difficult for the EU10 and the accession countries to promote tailored economic and social policies to suit there own particular needs. In particular we have argued that the type of institutions and policies that served Germany, Japan and South Korea so well are unlikely to be available. Both the macroeconomic and financial aspects of the dominant EU approach ignore the needs of the real economy, or at least show a remarkable confidence, not borne out by recent economic performance, in the ability of market forces to promote economic stability, growth and cohesion. Our reading is, therefore, that there is a pressing need for significant changes in the priorities, policies and institutions *within* the EU. The EU cannot be ignored, in part because some of the institutions and economic policies supported by economists in the Keynesian and post-Keynesian tradition are difficult to implement by small economies by their own. Tony Thirlwall, on the other hand, has been a vociferous critic of further European integration. Whatever the merits of this argument, we both agree that narrowing the options is unlikely to lead to enhanced economic performance let alone greater social cohesion.

NOTES

Acknowledgements: A much earlier version of this chapter was given at a FISC conference 'Global Finance and Social Europe', London, November 2003. We are grateful for the comments we received. The arguments presented in the chapter reflect the views of the authors themselves and should not be ascribed to the institutions they represent.

1. Tony Thirlwall, as all who know him will attest, has never been short of strong opinions. But he is the type of academic who, while vigorously defending his corner when amongst his peers, has always shown infinite kindness and tolerance to post-graduate students and younger academics. Since both the authors started their careers as lecturers at the University of Kent, it is appropriate to acknowledge here that we are both the beneficiaries of that kindness and tolerance. We shall always be grateful for his helping hand.
2. For an account of institutional economies and an assessment of their advantages and disadvantages with respect to more liberal economies see Crouch and Streeck (1997) and Hollingsworth and Boyer (1997).
3. The characteristics of ERM II are discussed in more detail in Gibson and Tsakalotos (2002).

4. See Dominguez and Frankel (1993). It is also thought that the upheavals in the old ERM in 1992 and 1993 were not helped by the stance of the Bundesbank on intervention to support affected currencies (Cobham, 1994).
5. As Keynes warned in 1941 'Loose funds may sweep around the world disorganising all steady business. Nothing is more certain than that the movement of capital funds must be regulated; – which in itself will involve far-reaching departures from laissez-faire arrangements' (Quoted in Thirlwall, 1999).
6. For an excellent critique of this literature, both on the underlying theory and the empirical evidence see Forder (2000).
7. See, for instance, Giavazzi and Pagano (1990).
8. Stiglitz (2004) argues that the 1990s boom in the US cannot be reduced to Clinton's fiscal consolidation.
9. Skidelsky (1992, chapter 11) provides a fascinating account of the economic arguments during this period and draws our attention to the extent to which modern concerns with credibility, and even Ricardian equivalence, have long histories.
10. As Keynes argued during the Great Depression, 'At the present time all governments have large deficits. [They are] nature's remedy, so to speak, for preventing business losses from being … so great as to bring production altogether to a standstill' (Quoted in Skidelsky, 1992, p. 394). Fitoussi (1997) argues that European policy-makers failed to take this into account in the late 80s and early 90s, and reacted to rising budget deficits with an overall stance that was unnecessarily restrictive.
11. Writing in 1931 Keynes believed that while inflation can be a problem 'it is worse in an impoverished world to provoke unemployment than to disappoint the rentier' (Quoted in Thirlwall, 1999). Henley and Tsakalotos (1993b) argue that Keynes optimism concerning inflation control was related to his belief that a certain kind of elite could be trusted not to act irresponsibly. In a modern age this rather elitist conception is unlikely to cut much ice, and in this context one must either rely on unemployment or corporatist arrangements for the problem of inflation – we return to the issue of corporatism below.
12. Whilst debt tends to be low in acceding countries, budget deficits have proved more difficult to control on a sustained basis. Interestingly, given our concerns in the second section of this chapter, one way of controlling debt is by using privatisation receipts. However, privatisation generates more of an equity culture and encourages the entry of Anglo-Saxon style institutional investors with the result that outsider financial systems become prevalent.
13. Moreover Forder argues that the conception of democracy that credibility theorists seem to think is not threatened, by such institutions as an independent central bank, is a very limited one. In particular he argues that this conception is restricted to instrumentalist accounts of the value of democracy.
14. Indeed, in official documents, emphasis is always placed on the independence of the various parties and the fact that that independence should not in any way be compromised (European Council, *Presidency Conclusions*, Cologne, pp. 5 and 8).
15. Thus, in the BEPGs, there is an emphasis on the need for wage agreements not to imitate each other across countries and to take into account regional differences in productivity growth, skill levels and unemployment. The decentralising dynamic behind this is that the EU is promoting the idea that these differences are not just between countries within EMU but within regions themselves in the different countries.
16. On Kaldor's view that monetarism entailed the deliberate use of unemployment to discipline labour see Thirlwall (1994). Henley and Tsakalotos (1993) discuss the role of social corporatism as an alternative to the deliberate increase in unemployment.
17. In part this approach has been inspired by the New Keynesian approach to the labour market. For a discussion of the theoretical underpinnings of the European Employment Strategy see Tsakalotos (2001).
18. See Henley and Tsakalotos (1993).
19. We discuss these issues in more depth in Gibson and Tsakalotos (2003).
20. Documentation relating to the Financial Services Action Plan is available at http://europa.eu.int/comm/internal_market/en/finances/actionplan/. For the Risk Capital Action Plan go to http://europa.eu.int/comm/economy_finance/publications/riskcapital_en.htm.

21. The Take-over Bids Directive initially was voted down by the European Parliament. A new and much watered down version was eventually adopted in April 2004.
22. For an up to date survey of the literature see Gibson and Tsakalotos (2003).
23. For a critique of the evidence on the supposed superiority of liberal finance institutions see Zhu, Ash and Pollin (2002).
24. It is worth noting that Anglo-Saxon style financial systems are not know for their generosity in providing funds either for small or medium-sized enterprises (SMEs, because the institutional investors which dominate financial markets find larger companies easier to follow) or for manufacturing companies (because they tend to be capital intensive and do not necessarily generate a lot of cash). Given that growth dynamics in the EU10 are heavily dependent on SMEs and that manufacturing may be especially important as an engine of growth at lower levels of development, the promotion of Anglo-Saxon style systems is even less appropriate.
25. Curiously Thirlwall (2000) ends up with a more upbeat portrayal of some of the supply-side aspects of the 80s economic policy, which it is argued were negated by poor macroeconomic policy. But this, as far as we can ascertain is not typical.

REFERENCES

Allen, F. and Gale, D. (2000), *Comparing Financial Systems*, Cambridge, MA: MIT Press.

Camerer, C. (1989), 'Bubbles and fads in asset prices', *Journal of Economic Surveys*, **3**, 3–41.

Chang, Ha-Joon (2002), *Kicking Away the Ladder: Development Strategy in Historical Perspective*, London: Anthem.

Cobham, D. (1994), *European Monetary Upheavals*, Manchester: Manchester University Press.

Crouch, C. (1997), 'The terms of the neoliberal consensus', *Political Quarterly*, **68**(4), 352–60.

Crouch, C. and Streeck, W. (eds) (1997), *Political Economy of Modern Capitalism*, London: Sage.

De Grauwe, P. (2004), 'The EU could use some economic optimism', *The Guardian*, 9 June 2004.

Dickerson, A.P., Gibson, H.D. and Tsakalotos, E. (1995), 'Short-termism and underinvestment: the influence of financial systems', *The Manchester School*, **LXIII**, 351–67.

Dominguez, K.M. and Frankel, J.A. (1993), *Does Foreign Exchange Intervention Work?*, Washington, DC: Institute for International Economics.

European Commission (2003), *2000 Broad Guidelines of the Economic Policies of the Member States and the Community*, Commission Recommendation, Brussels, available from www.europa.eu.int/eur-lex/pri/en/lip/latest/doc/2003/com2003_0170en02.doc.

European Council (various years), *Presidency Conclusions*, Brussels, available from www.europa.eu.int/council/off/counclu.

European Council (1998), *The Broad Economic Policy Guidelines*, Brussels, available from www.ue.eu.int/emu/broad.

European Council (1999), *1999 Broad Economic Policy Guidelines*, Brussels, published in *European Economy*, no.68.

Fitoussi, J-P. (1997) 'Is restrictive macro-economic policy the only way to EMU?', mimeo, April.

Forder, J. (1999), *Both Sides of the Coin: A Debate on the Euro*, London: Profile Books.

Forder, J. (2000), 'The theory of credibility: confusions, limitations, and dangers', *International Papers in Political Economy*, **7**(2).
Giavazzi, F. and Pagano, M. (1990), 'Can severe fiscal contractions be expansionary? Tales for two small European countries' in O.J. Blanchard and S. Fischer (eds), *NBER Macroeconomics Annual*, Cambridge MA: MIT Press.
Gibson, H.D. and Tsakalotos, E. (1994), 'The Scope and Limits of Financial Liberalization: a critical survey', *Journal of Development Studies*, **30**, 578–628.
Gibson, H.D. and Tsakalotos, E. (2002), 'EU enlargement, ERM II and the lessons from the southern European countries', mimeo, November.
Gibson, H.D. and Tsakalotos, E. (2003), 'Finance-industry relationships in Europe and the prospects for growth and convergence', in P. Arestis, M. Baddeley and J. McCombie (eds), *Globalisation, Regionalism and Economic Activity*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Gibson, H.D. and Tsakalotos, E. (2004), 'Capital flows and speculative attacks in prospective EU member states', *Economics of Transition*, **12**(3), September, 559–86.
Grahl, J. (2001), 'Globalized finance: the challenge to the Euro', *New Left Review*, **8** (new series), 23–47.
Henley, A. and Tsakalotos, E. (1993), *Corporatism and Economic Performance: a Comparative Analysis of Market Economies*, Aldershot, UK and Brookfield, USA: Edward Elgar.
Henley, A. and Tsakalotos, E. (1993b), 'Monetary rules versus consensual discretion: corporatism and the future of Keynesian policy making', in D. Crabtree and A.P. Thirlwall (eds), *Keynes and the Role of the State*, London: Macmillan.
Hollingsworth, J. and Boyer, R. (eds) (1997), *Contemporary Capitalism: the Embeddedness of Institutions*, Cambridge: Cambridge University Press.
Kay, J. (1993) *The Foundations of Corporate Success*, Oxford: Oxford University Press.
McCombie, J. and Thirlwall, A.P. (1992), 'The re-emergence of the balance of payments constraint', in Michie, J. (ed.), *The Economic Legacy 1979–1992*, London: Academic Press.
Mayer, C. (1994), 'The assessment: money and banking: theory and evidence', *Oxford Review of Economic Policy*, **10**, 1–13.
Modigliani, F., Fitoussi, J-P., Moro, B., Snower, D., Solow, R., Steinherr, A. and Labini, P.S. (1998), 'An economists' manifesto on unemployment in the European Union', *Banca Nazionale del Lavoro Quarterly Review*, **206**, September, 1–35.
Petersen, M.A. and Rajan, R.G. (1994), 'The benefits of lending relationships: evidence from small business data', *Journal of Finance*, **XLIX**, 3–37.
Shiller, R.J. (1981), 'Do stock prices move too much to be justified by subsequent changes in dividends?', *American Economic Review*, **71**, 421–36.
Shleifer, A. and Summers, L. (1988), 'Breach of trust in hostile takeovers' in A.J. Auerbach (ed.), *Corporate Takeovers: Causes and Consequences*, Chicago: NBER, Chicago University Press.
Skidelsky, R. (1992), *John Maynard Keynes: the Economist as Saviour 1920–1937*, London: Macmillan.
Streeck, W. (1997), 'Beneficial constraints: on the economic limits of rational voluntarism', in J. Hollingsworth and R. Boyer (eds) (1997).
Stiglitz, J.E. (1989), 'Financial markets and development', *Oxford Review of Economic Policy*, **5**, 55–67.
Stiglitz, J.E. (2004), 'The parties' flip-flops on deficit spending: economics or politics?', *The Economists' Voice*, **1**(1).

Teague, P. (1999), 'Reshaping employment regimes in Europe: policy shifts alongside boundary changes', *Journal of Public Policy*, **19**(1), 33–62.
Thirlwall, A.P. (1994), 'Nicholas Kaldor', in G. Hodgson et al. (eds), *The Elgar Companion to Institutional and Evolutionary Economics*, Aldershot, UK and Brookfield, USA: Edward Elgar.
Thirlwall, A.P. (1999), 'A second edition of Keynes' *General Theory*', *Journal of Post-Keynesian Economics*, Spring.
Thirlwall, A.P. (2000), *The Euro and 'Regional Divergence' in Europe*, London: New Europe Research Trust.
Thirlwall, A.P. (2001), 'Inflation is no devil in monetary policy detail', *The Guardian*, 24 December.
Thirlwall, A.P. (2003), *Trade, the Balance of Payments and Exchange Rate Policy in Developing Countries*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Tobin, J. (1984), 'On the efficiency of the financial system', *Lloyds Bank Review*, **153**, 1–15.
Tsakalotos, E. (2001), 'European employment policies: a new social democratic model for Europe?', in P. Arestis and M. Sawyer (eds), *The Economics of the Third Way*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
Zhu, A., Ash, M. and Pollin, R. (2002), 'Stock market liquidity and economic growth: a critical appraisal of the Levine/Zervos Model', *Working Paper Series*, no. 47, Political Economy Research Institute, University of Massachusetts.

12. Competition and competition policy in emerging markets: international and developmental dimensions*

Ajit Singh†

1. INTRODUCTION – COMPETITION POLICY AND DEVELOPING COUNTRIES: THE INTERNATIONAL CONTEXT

Developing countries are today faced with a range of new issues related to the microeconomic behaviour of economic agents – individuals, households and corporations – in these societies. In the past, such behaviour, and a country's institutional arrangements which supported it, has been the prerogative of sovereign nation states. However, with liberalisation and globalisation these matters are today regarded as legitimate objects of attention by the international community. Hence, under the new International Financial Architecture which is being constructed following the Asian crisis, emerging countries are being asked to reform their systems of corporate governance, labour laws, competition policy and other similar institutional structures. With respect to competition policy, which is the subject of this chapter, it is suggested by many policy makers that not only do developing countries require a competition policy, but a multilateral one would be greatly to their advantage.

The main purpose of this chapter is to brief developing countries on the complexities of this issue as well as its important policy implications for economic development. The chapter will examine the virtues and limitations of both national and international competition policies.

Contrary to the wishes of developing countries, the so-called "Singapore issues" were included in the WTO's November 2001 Doha Declaration of Ministers: these are investment, competition policy, trade facilitation and government procurement.[1] Competition policy was put on the agenda at the Singapore Ministerial meeting in 1996 as part of a review of the relationship between trade and investment. As this topic was being included in the WTO's work program – even at that time over the objections of developing countries – it was agreed that the matter should be studied by a working group with a remit

to pay particular attention to the development dimension of competition policy. This was to be without prejudice to the question of any prospective negotiations on the subject.[2] However, five years later at Doha, in one of the more confusing paragraphs of the Declaration, Ministers "*agreed that negotiations will take place after the fifth Session of the Ministerial Conference on the basis of a decision to be taken, by explicit consensus, at that Session on modalities of negotiations*". Many, but by no means all,[3] developed countries consider this as a mandate to launch negotiations at the fifth Ministerial in 2003 or shortly thereafter, whereas most developing countries maintain that the negotiations may be years off, as a decision to launch them must be taken by "explicit consensus". Much of this divergence arises from the undefined word "modalities", which countries choose to interpret in different ways.[4]

At India's request, Yussef Hussain Kamal, the Conference Chair at Doha, presented the following clarification: "In my view, this would give each Member the right to take a position on modalities that would prevent negotiations from proceeding after the fifth Session of the Ministerial Conference until that Member is prepared to join in an explicit consensus" (ICTSD, 2001, p. 5). As the clarification seems to express only a personal view, the legal status of the Chair's statement remains unclear. It is not formally attached to the Ministerial Declaration itself, but forms part of the official Conference proceedings.

Be that as it may, it is quite clear that sooner or later developing countries will need to be ready to enter into discussions or negotiations with advanced countries with respect to competition policy at the WTO as well as other multilateral, regional or bilateral fora.[5] International concern about the state of competition and competition policy in emerging countries precedes and goes beyond the Doha Declaration. This is because these issues also derive their international significance from some important analyses of the Asian financial crisis of 1997–98 and the subsequent proposals on the New International Financial Architecture. Competition and competition policy figure prominently in these designs for a new architecture for the global economic system. This is due to the fact that international financial institutions and orthodox economists suggest that the "deeper causes" of the recent Asian crisis were not the observed macroeconomic disequilibria but rather structural, linked to the normal Asian way of doing business. Apart from crony capitalism and close relationships between firms, banks and governments, such analyses single out for particular attention the allegedly poor competitive environment in the crisis-affected countries (Thailand, Indonesia and Korea). Further, in order to forestall future crises, it is argued that emerging markets need to be more open, transparent and "competitive".[6]

Nevertheless, it will be emphasised here that apart from these international dimensions, competition and competition policy are also important for developing countries in their own right. The present chapter builds on the author's

previous work in this area (Singh and Dhumale, 1999; Singh 2001a, 2001b) and extends it in a number of directions including specifically the analysis of:

(1) the relationship between competition, competition policy and development at the national level;
(2) the important implications of the recent new conceptual advances in the theory of industrial organisation for competition policy in developed and developing countries;
(3) the impact of market power exercised by industrial country firms on developing countries, including a more complete examination of the effects of the current cross-border international merger wave;
(4) in addition, the chapter puts forward a proposal for establishing an international competition authority to monitor anti-competitive behaviour by large multinationals, and discusses the desirability of such an authority, and what form it should take to address the concerns and particular needs of developing countries.

The chapter is organised as follows. Section II will consider the current state of competition and national competition policies in emerging markets and examine the relationship between competition, competition policy and economic development. Section III argues that although developing countries may not have needed competition policies in the past, they do so now in the wake of liberalisation and globalisation and the structural changes that these have brought about both at the national level (privatisation and deregulation) and at the international level (the gigantic international merger movement of the 1990s). Section IV will examine competition policy in the United States, the European Community and Japan in the light of new developments in economic theory and draw implications for developing countries. Sections III and IV are concerned, by and large, with national competition policies. Section V considers the desirability or otherwise of multilateral competition policy for developing countries. In that context it also examines a proposal for an international competition authority. Section VI concludes by summing up the main message of the chapter.

2. COMPETITION AND COMPETITION POLICY

2.a. Competition and Competition Policy in Emerging Markets

2.a.1. The state of competition in emerging markets

What is the nature of competition in emerging markets and how intensive is it? Strange as it may seem, in the light of market-oriented reforms which many

developing countries have been implementing over the last two decades, there are not many empirical studies on this topic.

There are a bare handful of comparative international studies for developing countries which provide data on variables such as three-or four-firm concentration ratios. Even this information tends to be somewhat dated. There also exist for a few countries more detailed studies usually in the standard structure-conduct-performance (SCP) paradigm.

In the absence of hard evidence, it is not surprising that there is considerable disagreement among economists speculating about the degree of competition in developing countries.

Laffont (1998) suggests that these countries exhibit segmented product markets, discretionary government regulations and considerable corruption and hence are not very competitive. As noted earlier, the advocates of the structuralist theory of the Asian financial crisis of 1997–98 believe that the crisis-affected Asian countries, including Korea, suffered from poor competitive environments that resulted in overinvestment. Michael Porter (1990), on the other hand, suggests that the Korean *chaebol* (large conglomerates) display highly rivalrous behaviour, and in the areas where Korea has been internationally successful, these companies have been subject to intense national and international competition.

Some apparent support for the lack of competition referred to above is provided by evidence on how relatively difficult it is to start a new business in emerging markets, due to complex government regulations and bureaucracy.[7] There are also considerable barriers to exit in many developing countries. Further, there is evidence that many developing countries favour large firms at the expense of small firms in the provision of finance and other measures.

Data in Table 12.1 on concentration ratios lend some support to the competition deficit thesis. The table indicates that concentration ratios in developing countries have been quite high relative to advanced countries. However, Table 12.2, which reports the share of small enterprises in total employment, suggests the opposite, i.e. that there may be greater competition in developing than in advanced countries. The differences between the US and the developing countries in Table 12.2 are quite dramatic. Whereas small enterprises (accounting for less than 10 workers) account for about 4% of total employment in the US, in emerging countries the share is several orders of magnitude higher. The data in Table 12.2 are subject to some important statistical biases, all of which would, if anything, understate the share of small firms in the economies of developing countries.[8] In more general terms, what the data in tables 12.1 and 12.2 indicate is the dualistic structure of developing country economies: a large modern sector accounting for a big proportion of total output exists side by side with a very large traditional sector of small enterprises, which contribute an almost equal proportion of the economy's output. Relative to advanced countries, the share

Table 12.1 Concentration ratios in emerging markets

Economy Share	**Three-firm concentration ratios**
Japan, 1980	56
Korea, Rep. of, 1981	62
Taiwan, China, 1981	49
	Four-firm concentration ratios
Argentina, 1984	43
Brazil, 1980	51
Chile, 1979	50
India, 1984	46
Indonesia, 1985	56
Mexico, 1980	48
Pakistan, 1985	68
Turkey, 1976	67
United States, 1972	***40***

Source: World Bank (1993).

of the small-scale sector in developing countries in terms of employment would be larger than in terms of output because of the bigger differences in capital intensity of the two sectors in these countries.

Thus, as far as the intensity of competition in the two groups of countries is concerned, Tables 12.1 and 12.2 provide conflicting information. Moreover, these data bear only on the static measures of concentration, which have acknowledged shortcomings as indicators of the intensity of competition. To obtain a more complete picture of the competition process, it is necessary to supplement these static measures with indicators of the dynamics of the competition process. Fortunately there now exist some studies on this subject, and it is also useful that they employ different methodologies to model the dynamics. First, there is research by Glen, Lee and Singh (2001), which examines the persistency of profits in seven emerging markets in the 1980s and the early 1990s. The authors use exactly the same methodology as that employed by "persistence of profitability studies" for advanced countries. The results of their time series estimates of persistence coefficients in emerging markets are reported in Table 12.3. For purposes of comparison, Table 12.4 summarises the results of similar studies for advanced countries. Surprisingly, the results indicate that developing countries have consistently lower persistency coefficients than those observed for advanced countries, indicating that on the normal interpretation of such

Table 12.2 Distribution of employment shares for small enterprises in developing countries and the US

Number of Workers	1–4	5–9
United States, 1992	1.3	2.6
Mexico, 1993	13.8	4.5
Indonesia, 1986	44.2	
S. Korea, 1973	7.9	
S. Korea, 1988		12
Taiwan, 1986		20
India, 1971	42	
Tanzania, 1967	56	
Ghana, 1970	84	
Kenya, 1969	49	
Sierra Leone, 1974	90	
Indonesia, 1977	77	
Zambia, 1985	83	
Honduras, 1979	68	
Thailand, 1978	58	
Philippines, 1974	66	
Nigeria, 1972	59	
Jamaica, 1978	35	
Colombia, 1973	52	
Korea, 1975	40	

Source: Adapted from Tybout (2000).

Table 12.3 Developing countries: mean values of λ_i and proportion of significantly positive and significantly negative Y_{iLR}

	Mean λ_i	**Positive Y_{iLR}**	**Negative Y_{iLR}**
Brazil	0.013	1/56	3/56
India	0.229	2/40	4/40
Jordan	0.348	1/17	0/17
Korea	0.323	7/82	2/82
Malaysia	0.349	4/62	7/62
Mexico	0.222	0/39	0/39
Zimbabwe	0.421	0/40	4/40

Source: Glen, Lee and Singh (2002).

Table 12.4 Persistence of profitability studies for industrial countries

Author	Country	Sample Period	Observations per firm	Number of firms	Sample mean (Lamda [i])
Geroski and Jacquemin (1988)	UK	1947–77	29	51	0.488
	France	1965–82	18	55	0.412
	Germany	1961–81	21	28	0.410
Schwalbach et al. (1989)	Germany	1961–82	22	299	0.485
Mueller (1990)	US	1950–72	23	551	0.183
Cubbin and Geroski (1990)	UK	1948–77	30	243	0.482
Khemani and Shapiro (1990)	Canada	1964–82	19	129	0.425
Odagiri and Yamawaki (1990)	Japan	1964–82	19	376	0.465
Schohl (1990)	Germany	1961–81	21	283	0.509
Waring (1996)[c]	US	1970–89	20	12,986	0.540

Source: Goddard and Wilson (1999).

results, developing countries are subject to no less, if not greater competition, than advanced countries. The possible sources of bias in the empirical results for emerging economies have been examined by Glen, Lee and Singh (2002) and they find that these do not affect their main conclusions.

Complementary evidence to that of Glen, Lee and Singh is provided by another kind of research, which also bears on the dynamics of the competition process but uses a different methodology. This research, which examines turnover, the entry and exit of firms, provides extremely interesting results. Some of the latter are summarised in Table 12.5. The Table indicates that there is greater turnover as well as entry and exit of firms in the small number of emerging markets for which such studies have been carried out than for advanced countries.

Apart from these two kinds of studies done on the dynamics of the competition process, there are also other types of evidence pertaining to the efficiency of emerging market industries and to scale economies that do not accord with the conventional anecdotal account of the lack of competition in emerging countries. This empirical research has recently been reviewed by Tybout (2000, p. 38) who sums up the situation as follows:

> Indeed, although the issue remains open, the existing empirical literature does not support the notion that LDC manufacturers are relatively stagnant and inefficient.

*Table 12.5 Plant and job turnover in developing versus developed countries**

	Turnover Rates				Market Share of Entrants			
	Plants		Jobs		< 1 year olds	< 5 year olds	< 10 year olds	Minimum Plant Size Covered
Country (period covered)	1 year	5 year	1 year	5 year				
Chile (1979–86)	8.5		26.9		3.6	15.3		10 workers
Colombia (1977–89)	11.9		24.6		4.9	19.8		10 workers
Morocco (1984–90)	9.5		30.7		3.2			10 workers
Korea (1983–93)		64.2				32.5		5 workers
Taiwan (1981–91)		67.9				43.9	63.2	1 worker[c]
United States (1963–82)		26.9[a]	18.9[b]	58.4		10.7[a]	18.6[a]	5 workers
Canada (1973–92)			21.9[b]					5 workers

Notes:

* Let N_t be the number of plants observed in year ,; E_t be the number of plants observed in year t but not $t-1$; and X_t be the number of plants observed in year $t-1$ but not in year t. Then the entry rate is E_t/N_{t-1} and the exit rate is X_t/N_{t-1}. The plant turnover rate is the average of these two statistics. Similarly, the rate of gross job creation is the number of jobs at entering plants plus the number of new jobs at expanding plants, divided by the initial number of jobs, and the gross job destruction rate is the number of jobs that disappear as plants contract or exit divided by the initial number of jobs. The sum of these two rates is the job turnover rate.

[a] Figures are average rates of *new* entry rates across 4-digit industries.

[b] These figures are for 1973–1992.

[c] The Taiwanese data set describes firms rather than plants.

Source: Tybout (2000), p. 26.

Turnover rates in plants and jobs are at least as high as those found in the OECD, and the amount of cross-plant dispersion in measured productivity rates is not generally greater. Also, although small-scale production is relatively common in LDCs, there do not appear to be major potential gains from better exploitation of scale economies.

2.a.2. The state of competition policy in developing countries

Most developing countries have, until recently, operated without a formal competition policy.[9] As Table 12.6 suggests, until 1990 only 16 developing countries had a formal competition policy. With encouragement and technical assistance from international financial institutions and the WTO, 50 countries have completed legislation for competition laws in the 1990s, and another 27 are in the process of doing so. It should, however, be borne in mind that it takes about ten years for countries to acquire the necessary expertise and experience to implement such laws effectively (Scherer, 1994).

Table 12.6 Number of developing countries that have adopted competition laws, as of June 2000

	Pre-1950s	1950s	1960s	1970s	1980s	1990s	Under preparation	Total
Asia / Pacific	0	0	2	2	2	14	6	26
Central / Eastern Europe	0	0	0	0	1	16	1	18
Latin America and Caribbean	1	2	1	1	0	6	10	21
Africa	0	1	0	1	2	14	10	28
Total	1	3	3	4	5	50	27	93

Source: Calculated from Table V.1, p. 151 of UNCTAD (2000).

The main reason why developing countries did not have a formal competition policy was that it was not needed. This is because there was considerable state control over economic activity, and if the government thought there was anti-competitive behaviour by some corporations or industries it intervened directly and fixed prices (such as for medicines and other essential products). Besides, state-owned industry was enjoined not to charge monopoly prices.

There is also evidence that competition laws have varied widely in countries where they have been introduced. Based on a survey of competition laws in fifty countries, World Bank (2002) reports that there are important inter-country differences in three dimensions: (a) the definition of dominance; (b) the treatment of cartels; and (c) enforcement. With respect to the definition of market

Table 12.7 Benchmarks of product market dominance in competition laws around the world

Country group	Market share of the firm
Developing and transition countries	
East Asia	50–75 percent
Eastern Europe and Central Asia	30–40 percent
Africa	20–45 percent
Industrial countries	
United States	Two-thirds or more
European Union	40–50 percent

Source: World Bank (2002), p. 140.

dominance, for example, a majority of countries define it in qualitative terms. However, 22 countries out of 50 define it quantitatively, although with widely varying thresholds, as seen in Table 12.7. Similarly, the treatment of cartels varies greatly in its severity. On the effective implementation of competition laws, the *World Competitiveness Yearbook, 2000*, estimates that, based on a survey of top-and middle-management of firms in each country studied, competition authorities in advanced countries are 40% more effective than their counterparts in developing countries.[10]

2.b. Competition, Competition Policy and Economic Development

2.b.1. Competition and competition policy

There would appear to be no obvious relationship between competition policy and competition, since, as we have seen, many developing countries have been able to maintain considerable competition in product markets despite the absence of a formal competition policy. An analytical reason for this lack of correspondence between competition and competition policy lies in the fact that developing country firms have been increasingly subject to foreign competition with the liberalisation of their economies. In the Asian NICs, even while they maintained selective import controls, their export-oriented policies exposed firms to competition in foreign markets. An additional reason was that governments in these countries organised contest-based competitions for state subsidies which were conditioned on the achievement of certain performance standards (export targets, foreign exchange earnings, and technological upgrading are a few of the contest objectives utilised by governments), with the winners receiving greater aid from the government.[11]

2.b.2. Competition and economic development

The relationship between competition and economic development is controversial, both in economic theory and in relation to empirical evidence. Economic orthodoxy posits a monotonic positive relationship between the two variables and therefore suggests that the greater the intensity of competition the better the economic performance.[12] However, modern economic analysis seriously qualifies that conclusion. As Telser (1987) observed, despite the reluctance of "many economists to accept the proposition that competition may be excessive because the received theory regards competition as always good, the more there is the better", new developments in the theory of industrial organisation indicate that the excess competition proposition is valid. These developments suggest that maximum competition is not necessarily the optimal degree of competition for promoting either economic welfare in the static sense, or, more importantly, in the dynamic sense, maximising the long-term trend-rate of growth of productivity in the economy.[13]

In the real world, it is recognised that the case for competition necessarily spurring economic efficiency at the micro-economic level is very weak because of the separation of management and control in large corporations, asymmetric information, transactions costs and agency problems. Indeed, Nickell (1996) suggests that the case for a positive link between competition and increased effort by economic agents is both theoretically tenuous and has little empirical support.[14] Nickell argues, therefore, that the virtues of competition are more convincing at the broad-brush impressionistic level rather than on the basis of rigorous econometric studies. He cites the broad, long-term experience of Japan (good – due to a high level of competition) and that of communist Eastern Europe (bad – because of lack of competition) as the best confirmation of the positive relationship between competition and economic development.

The seminal World Bank (1991) *Report* which provided the intellectual basis for the Washington Consensus contended on its first page that:

> Competitive markets are the best way yet found for efficiently organizing the production and distribution of goods and services. Domestic and external competition provides the incentives that unleash entrepreneurship and technological progress. (p. 1)

The *Report*'s basic analytical approach was to suggest that the fastest growing countries were those with the most rapid growth of total factor productivity (TFP). The latter in turn depended on domestic and foreign competition achieved through free markets. The role of the state was, in this view, essentially that of a "night watchman" concerned with providing human capital and physical and social infrastructure that provides a conducive environment in which business can flourish. This may be a plausible model in theory but in practice it did not describe the East Asian experience at all accurately. However, as World Bank

(1993) itself acknowledged in its subsequent report on the East Asian miracle, these countries did not have maximum competition in product, capital or labour markets, but rather strived to achieve an optimal degree of co-operation and competition. Thus, for example, the Japanese and Koreans implemented selective import controls; fostered close relationships between government, business and finance; and discouraged foreign investment while importing technology from abroad by other means.[15] The "broad-brush" East Asian evidence, in short, does not bear out Nickell's claims for the virtues of competition in relation to economic development (see Singh, 2003b). The experience of China, which for the last two decades has had one of the fastest growth rates in the world, is also consistent with this East Asian story. The Chinese economy has been able to register such fast growth rates despite its segmented product markets and highly imperfect capital and labour markets.

Nickell's (1996) own study reports a positive relationship between competition and long-run productivity growth for UK firms. He notes, however, that in general, empirical evidence for the claim that competition enhances corporate performance is not overwhelming. Detailed microeconomic research also indicates that there is no monotonic relationship between competition (as proxied by the number of firms) and managerial effort or other benefits of competition. On the basis of game-theoretical models as well as empirical studies, World Bank (2002, p. 134) notes that it is possible to attain the benefits of competition – greater efficiency and innovation in product markets – with "some" degree of competition, but competition by a large number of firms is not always required.

Another useful piece of evidence comes from an interesting recent study by Aw, Chung and Roberts (2002) that systematically compares turnover and exit and entry rates for South Korean and Taiwanese firms in seven comparable industries in the late 1980s. The results indicate that, on all the dynamic measures of competition examined by the authors, Taiwan was more competitive than South Korea. Nevertheless, it is worth noting that the overall trend rates of growth in the two economies have been very similar in the period examined as well as over a longer period.

To sum up, the main point that emerges from the above discussion of the relationship between competition and economic development is that a suitable combination of co-operation and competition is more likely to enhance societal welfare than competition alone. This conclusion is supported not only by the experience of the East Asia countries and China, but also by that of industrial districts in Italy and in many other countries.[15] Further, recent theoretical developments suggest that, in relation to innovation, "inter-firm co-ordination even among horizontal competitors can bring substantial welfare benefits." (Baumol, 2001 p. 736).

3. COMPETITION POLICY, LIBERALISATION AND GLOBALISATION

3.a. Competition Policy, Privatisation and Deregulation

Notwithstanding developing countries' lack of experience with competition policies, and the general scepticism about whether maximum or perfect competition is optimal for long-term productivity growth, there are good reasons to suggest that, under the present global economic arrangements, it is important for developing countries to establish formal competition policies. This is primarily because enormous structural changes have occurred in developing country economies during the last two decades as a result of privatisation and deregulation. These have been spawned by technological, economic, political and ideological forces, which are leading to liberalisation as well as greater integration of the world economy. As many of the privatised firms include natural monopolies, it is important that an appropriate regulatory and competition policy framework be in place to ensure improved economic performance. In relation to the question of perfect versus optimum degree of competition, it will be appreciated that nuanced competition laws will be required to implement optimal competition.

A significant danger is that privatisation may simply lead to a substitution of public sector monopolies by private monopolies, which would arguably reduce social welfare, as unlike in the case of the public sector, the private sector is usually under no formal injunction to advance people's wellbeing. In addition, the experience of privatisation in the UK and many other countries suggests that it is not ownership itself which affects performance, but rather the external environment, particularly as regards to competition, which is the more important factor.[17] Hence, the need for an appropriate competition policy.

3.b. The International Merger Wave and Competition Policy in Developing Countries

Another extremely important reason for developing countries to have competition laws today derives from the huge international cross-border merger movement which has been re-shaping the world economy during the last decade. UNCTAD (2000, 2001) data shows that the global value of cross-border acquisitions has risen from about 0.5% of world GDP in the mid-1980s to well over 2% in 2000. As these international mergers, as well as those between large corporations within the developed countries themselves, are quite central to the policy proposals which will be put forward later, it is important to carefully review the stylised facts and what the vast literature on mergers has taught us about these phenomena.

3.b.1. The 1990s global merger wave in historical perspective

The first stylised fact about mergers is that these normally come in waves. Secondly, analyses suggest that each wave generally has the stamp of special factors; these normally lead to differing perceptions concerning asset values among economic agents which in turn encourages mergers (Gort, 1969). Among the largest recorded waves in the United States is that between about 1890 and 1905.[18] This wave – dominated by "mergers for monopoly" – saw the creation of giant US firms which subsequently dominated the industrial landscape for much of the 20th century. Ironically, it is thought that the Sherman Antitrust Act of 1890, by outlawing co-operation between firms, thereby encouraged mergers.[19] The 1920s wave was labelled by Stigler as "mergers for oligopoly". The wave of the 1960s was characterised by conglomerate mergers and that of the 1980s by the "bust-up" of the same conglomerate mergers and by leveraged buy-outs.

The 1990s merger movement, in contrast, has witnessed "size-increasing blow-up mergers" creating very large global players. This wave had its origin in new technology, globalisation and deregulation, factors which, not surprisingly, lead to dramatic disturbances in economic agents' perceptions of market valuations of firms, fuelling mergers. Many of these mergers are defensive in that, once one large player takes over another company, other players are obliged to follow suit, through defensive takeovers, in order to maintain their market share. Holmstrom and Kaplan (2001) argue that a distinguishing feature of the 1990s merger wave in the US was enormous changes in corporate governance. In their view, during the 1980s the capital market exercised an increasing influence on corporate performance through leveraged buy-outs and other hostile tender offers for firms. However, by the 1990s, managers appear to have internalised the virtues of maximising shareholder value as their main motivation, not least because they themselves benefited through stock options. In brief, the work of leading scholars in this area suggests that the 1990s takeover wave in the US and in the world economy has been motivated by firms trying to achieve domination and bigger size in a global market. This has been effected by offensive takeovers in the market for corporate control, even though their intent in terms of product market competition may have been defensive, e.g. maintaining market share in the face of takeovers by competing firms.

The US merger wave of 1990s appears to have come to an end with the deflation of the technology bubble on the stock market in the first half of 2000. However, once the stock market revives, there continue to exist a number of factors which could propel another merger movement. In previous merger waves most mergers were national, that is, within national boundaries. This was particularly striking in the merger wave of the 1960s which took place simultaneously in most leading industrial economies. However, despite these waves occurring at the same time in many countries, there were surprisingly few cross-border

takeovers (Singh, 1992). In that sense, the 1990s merger wave, with a huge component of cross-border takeovers, has been quite unique.

It is also important to note that most cross-border mergers take place among industrial countries themselves. They are closely linked with foreign direct investment (FDI). Indeed, most of the FDI among advanced countries nowadays occurs through this channel. However, the incidence of cross-border takeovers via FDI is much lower in developing than in advanced countries. The overall estimate of the size of cross-border takeovers in developing countries (see Table 12.8) is somewhat misleading as it is heavily influenced by China. The latter is the largest developing country recipient of FDI, but it generally permits only greenfield FDI. If China is excluded, cross-border takeovers constitute a far larger and growing part of FDI in other developing economies (UNCTAD 1999).

Table 12.8 Cross-border M&As:[a] sales and purchases, 1998–99 (US$ billions)

Region/economy	**Sales**		**Purchases**	
	1998	1999	1998	1999
Developed countries:	**445.1**	**644.6**	**511.4**	**677.3**
European Union	187.9	344.5	284.4	497.7
United States	209.5	233.0	137.4	112.4
Japan	4.0	15.9	1.3	9.8
Developing countries	**80.7**	**63.4**	**19.2**	**41.2**
Africa	0.7	0.6	0.2	0.4
Latin American and the Caribbean	63.9	37.2	12.6	24.9
Europe		0.3		
Asia	16.1	25.3	6.4	15.9
Pacific		0.1		
Central and Easter Europe[b]	***5.1***	***10.3***	***1.0***	***1.6***
World[c]	**531.6**	**720.1**	**531.6**	**720.1**

Notes:
[a]Cross-border M&As that result in the acquisition of more than 10% equity share.
[b]Includes the countries of the former Yugoslavia.
[c]Includes amounts which cannot be allocatied by region.

Source: UNCTAD (2000).

3.b.2. Benefits and costs of mergers

What are likely to be the benefits and costs of the 1990s merger wave to different groups of countries and to the global economy? In theory, takeovers can increase societal welfare through two separate channels. The first is the threat of takeovers, which may oblige inefficient firms to become more efficient; the second is through the actual takeovers themselves, which may lead to synergies between the activities of the acquired and acquiring firms. An obvious social cost of the second channel is the potential monopoly power of the merged firms.

It is significant that, although takeovers and mergers are central to the theory of the firm, to industrial organisation, privatisation and deregulation amongst other fields of study, there is no unified theory of mergers as such (see further Singh, 1992). However, issues concerning the costs and benefits of mergers have been explored in a vast literature comprising both analytical and empirical studies.[20] In order to give coherence to the empirical studies, I shall briefly review the work on the nature of the takeover selection mechanism on the stock market; this includes, *inter alia*, the question above concerning the extent to which the threat of takeovers is effective in improving economic performance of inefficient firms (measured by profitability). Singh 1971 and 1975, which were among the first studies on this subject, investigated this by comparing the multivariate characteristics of: (a) acquired and non-acquired firms; (b) acquiring and acquired firms; and (c) acquiring and non-acquiring firms. Briefly, Singh's main result was that selection in the market for corporate control does not occur solely on the basis of performance (e.g. profitability or stock market valuation); it also occurs on the basis of size, which is the more important discriminator. Thus, a large unprofitable firm has a better chance of survival than a small, relatively more profitable firm. These results on the empirical characterisation of the selection mechanism have been confirmed in many subsequent studies (see Table 12.9 which summarises a wide range of studies from different countries on this and other points related to takeovers).

The effects of mergers have been studied by industrial organisation economists in terms of profitability and by financial economists in terms of stock market valuation. Most studies by industrial organisation economists invariably find reduced profitability after mergers, or, at the best, no change, after controlling for all the relevant factors.[21] Financial economists, on the other hand, believe that mergers increase the stock market value of the combined firms. This valuation undoubtedly increases during a short period of a few weeks preceding the takeover event. At that time, the acquired firm's value increases by an average of 20–30%; the acquiring firm's value remains more or less the same. The combined result is greater value. However, the acquiring firm's shareholders suffer systematic losses which begin as soon as six months after the takeover and which may go on for a number of years.

Table 12.9 Merger studies results: a summary

Characteristics	Bidder	Target
Size	>industry, > target	< industry, < bidder
Profitability	> industry, > target	> industry, < bidder
Growth	> industry, > target	< industry, < bidder
Market-to-book ratio	< industry	?
Effects of the acquisition		
Return on share	announcement about equal to long-run losses premium 20–50 percent	
Profits	down	
Sales	down	
Market share	down, in most cases at the loss of the target	
Investment	about same	
R&D	about same	
Asset restructuring	yes	
Management turnover	yes	
Labour costs	down	
Elements of success		
Relatedness	related business > horizontal > vertical > conglomerate	
Size difference	big difference > equal size	
Motive	tender offer > merger; hostile ? Friendly	
Market-to-book ratio	value bidders > glamour bidders	
Financing	cash > stock	

Source: Tichy (2002).

On the question of whether mergers lead to concentration or monopoly power, there is a large and controversial literature. On one side are economists who believe that with liberalisation of trade throughout the world, the size of the relevant market has enormously increased and therefore the monopoly power effects of mergers are no longer significant. This view is strongly contested by others. Tichy (2002), in his comprehensive survey of the merger literature, notes that

> Contrary to widely held opinion, concentration is a quickly increasing problem, even with the extension of markets resulting from the new policy of the big corporations. Driven by the superficial advice of their consultants they strive hard to belong to the top three in their respective market, and sometimes they concentrate forcefully on narrowly limited markets to achieve this goal. If they are successful, a great number of oligopolistic markets with very few competitors result, strongly supporting collusion. (Tichy, 2002, p. 20)

He concludes by observing that the "goal of being among the three leading players in the world market creates oligopolistic power if the competitive fringe is not extremely strong, as the likelihood of strong competition with fewer than four to five competitors is rather small".

It will be useful at this point to review a recent massive study of the effects of mergers carried out by Gugler *et al*. (2003). They examined mergers of the decades of the 1980s and the 1990s in a large number of countries in all parts of the world, where the relevant data is available. Defining merger as a transaction where more than 50% of the equity of the "victim" firm is acquired, their data suggests that between 1981 and 1998 there were nearly 70,000 merger announcements; of these 45,000 were actually completed, nearly half of them in the United States (see Table 12.10). The results of their study of the effects of these mergers on profits, sales and market value and their overall findings are reported in Table 12.11. These results are broadly consistent with those of much of the merger literature. The effects of mergers on profitability are positive but insignificant, until the fifth year after merger when the positive effect is significant only at the 10% level. The impact on sales and market value are strongly negative and statistically significant from the merger year onwards.

In terms of the methodology presented by Gugler *et al.*, mergers which enhance the efficiency of the merging firms should demonstrate an increase in both their profits and their sales. An efficiency-reducing merger would have a negative effect both on sales and on profits. A merger which increases market power would increase profits and reduce sales. Overall, the authors' results indicate that no more than a quarter of the mergers appear to increase efficiency, and a quarter increase profits by increasing the market power of the firms involved. About 50% of the mergers fail, paying for increased sales by reduced profits or losing sales as well as profits. The authors of this study were also able to compare the effects of cross-border mergers with those of domestic ones. These results suggest broadly similar effects.

The above analysis of the effects of mergers has taken a rather narrow, economic efficiency view of their benefits and costs. A more comprehensive analysis must also include a discussion of the distributional consequences of mergers, as these often tend to be quite large. The benefits of mergers may, for example, go to shareholders whilst the costs may be borne by workers who lose their jobs as a result of rationalisation. Although the importance of these distributional issues is recognised (see for example Shleifer and Summers, 1988; Singh, 1993), there is very little empirical literature on the subject. In the context of the present chapter the issue of the distribution of the gains and costs of mergers between countries is also pertinent.

Table 12.10 Number of deals, value of deals (millions of dollars), percentage of cross-border deals, percentage of horizontal and conglomerate mergers by years and country groupings

	Year					
	until 1990	1991/92	1993/94	1995/96	1997/98	Total
US						
Number of deals	8194	1965	2840	3783	4369	21151
Value of deals	242.44	104.45	139.33	195.70	313.94	221.19
% cross-border	0.03	0.12	0.14	0.16	0.17	0.11
% horizontal	0.38	0.47	0.49	0.49	0.48	0.44
% conglomerate	0.56	0.48	0.47	0.48	0.49	0.51
UK						
Number of deals	1180	501	790	1138	1148	4757
Value of deals	220.07	113.82	61.11	97.13	158.92	139.32
% cross-border	0.35	0.30	0.27	0.27	0.36	0.32
% horizontal	0.31	0.36	0.35	0.38	0.36	0.35
% conglomerate	0.64	0.59	0.62	0.58	0.61	0.61
Western Europe						
Number of deals	986	2125	1996	2364	2059	9530
Value of deals	398.95	188.63	163.41	144.44	320.33	241.90
% cross-border	0.54	0.24	0.27	0.33	0.48	0.35
% horizontal	0.37	0.44	0.37	0.36	0.34	0.38
% conglomerate	0.59	0.53	0.59	0.61	0.63	0.59
Japan						
Number of deals	172	88	61	151	168	640
Value of deals	528.91	474.11	198.55	754.97	177.43	478.73
% cross-border	0.81	0.72	0.59	0.34	0.45	0.57
% horizontal	0.33	0.30	0.36	0.35	0.35	0.34
% conglomerate	0.62	0.70	0.61	0.63	0.61	0.63
Aus./NZ/Canada						
Number of deals	671	425	549	767	875	3287
Value of deals	357.63	69.55	61.56	126.97	109.70	150.54
% cross-border	0.38	0.23	0.32	0.28	0.36	0.32
% horizontal	0.44	0.43	0.48	0.40	0.39	0.42
% conglomerate	0.52	0.55	0.49	0.57	0.58	0.54
Rest of world						
Number of deals	371	553	831	1731	1733	5219
Value of deals	278.88	150.74	88.64	112.76	142.92	132.60
% cross-border	0.50	0.26	0.33	0.25	0.35	0.31
% horizontal	0.34	0.36	0.35	0.37	0.37	0.36
% conglomerate	0.59	0.59	0.63	0.60	0.60	0.60
Total						
Number of deals	11574	5657	7067	9934	10352	44584
Value of deals	260.63	131.11	116.51	162.7	243.09	199.71
% cross-border	0.16	0.21	0.23	0.24	0.30	0.23
% horizontal	0.37	0.43	0.42	0.42	0.41	0.41
% conglomerate	0.57	0.53	0.54	0.55	0.56	0.55

Source: Gugler *et al*. (2003).

Table 12.11 Effects of mergers for full sample

Years after the merger	***Profitability***				***Sales***				***Market value***			
	Difference (Mn. US$)	p-value	Obs.	Pos.	Difference (Mn. US$)	p-value	Obs.	Pos.	Difference (Mn. US$)	p-value	Obs.	Pos.
0	.	.	.	.	.	.	.	.	50.512	0.000	6211	0.524
1	1.884	0.580	2603	0.554	73.654	0.008	2603	0.688	–49.262	0.035	5282	0.472
2	3.864	0.413	2171	0.556	–87.435	0.025	2171	0.652	–177.464	0.000	4189	0.437
3	5.081	0.395	1810	0.536	–193.993	0.000	1810	0.621	–342.415	0.000	3416	0.401
4	11.660	0.119	1474	0.566	–189.038	0.008	1474	0.610	–528.251	0.000	2784	0.375
5	15.989	0.097	1210	0.565	–263.392	0.004	1210	0.588	–865.520	0.000	2218	0.357

Source: Gugler *et al.* (2003), p. 29.

3.b.3. The international merger movement and developing countries

The foregoing examination of the causes and effects of the current international merger movement raises important concerns for developing countries. First, there are the obvious questions of increased market power of large multinationals and their potential abuse of this dominance. Developing countries are clearly affected directly by the monopoly power effects of international mergers, as when a foreign multinational acquires a domestic firm. However, they are also affected indirectly even when mergers take place outside their jurisdictions, e.g. within advanced countries themselves. The "rule of being in the top three", as Tichy argues, reduces the contestability of these markets and is especially harmful to the interests of late industrialising countries whose firms are building up their capabilities to compete in international markets. The reduced contestability of markets is therefore of special concern for developing countries.

Developing countries clearly need a competition policy in order to be able to deal with these issues of market dominance and abuse of dominant positions. However, even with such legislation on the statute books these countries may not have the power to restrain cartels and other uncompetitive conduct by large multinationals, owing to inadequate development of their legal and institutional frameworks, lack of information and difficulties of proving that prices are being manipulated by international cartels. It has become conventional to underplay the practical significance of cartels, presumably on the ground that these arrangements tend to be short-lived and their incidence is quite low. However, the US anti-trust authorities, which have long held a strong anti-cartel position, made their stance even stronger in the early 1990s. By the end of the decade the US position was accepted by EU and other advanced countries. Consequently, several industrial countries have passed legislation to stiffen the penalties for participation in illegal cartels. It is also increasingly recognised that the illegal cartels that are actually detected and prosecuted are merely the tip of a large iceberg.[22] Recently, the US government fined participants in a European vitamins cartel a record sum of seven hundred and fifty million dollars. If such cartels can operate in an economy like that of the US, with its long history of anti-trust laws and their enforcement, it is more than likely that their incidence will be quite high in developing countries.

The experience of the 1930s with widespread cartelisation in peripheral countries points in the same direction. During that decade, it is estimated that very roughly 30 to 50% of the world's exports were subject to cartel manipulation of prices. The post-World War II reduction in the incidence of cartelisation may mainly be attributed to two factors: (a) stricter enforcement of laws against cartels in advanced countries, particularly the US; and (b) the replacement of cartels by straightforward mergers between firms.[23] However, more recently, the EU Competition Commission has begun to vigorously prosecute cartel arrangements in many different industries, so as to ensure that a single European

market is not thwarted by cartelisation and division of markets.[24] The latter was widespread in the 1930s in the so-called sphere of influence cartelisation. It often took the form of European companies, for example, withdrawing from competition with the US firms in Latin American countries as these lay in the American sphere of influence. In return, the US corporations would undertake not to compete with Europeans in Southern and Eastern Europe, recognising these countries to be in the European sphere of influence.

It is important to note that the anti-cartel legislation existing in advanced countries does not normally extend to developing countries. Indeed, on the contrary, exports or foreign markets are often explicitly exempted from such laws. In these circumstances, in addition to domestic competition policies, developing countries clearly require considerable co-operation from advanced countries to be able to cope at all effectively with the anti-competitive behaviour of advanced country cartels between the large multinationals. From the perspective of poor countries, it is therefore necessary not only to have the right kind of domestic competition policies, but also an appropriate framework for international cooperation on competition issues.

4. COMPETITION POLICY AND DEVELOPING COUNTRIES: TAKING ACCOUNT OF THE DEVELOPMENTAL DIMENSION

What kind of competition policy would be suitable for developing countries? In the light of the previous discussion, such a policy must at least be able to: (a) restrain anti-competitive behaviour by domestic privatised large firms; (b) limit abuses of monopoly power by mega-corporations created by the international merger movement; and (c) promote development.

The question of what constitutes an appropriate competition policy for late industrialising economies will be examined here on the basis of economic theory and the historical experience of the developed countries – the European Union countries, the United States and Japan.

4.a. Competition Policy and Development: Analytical Considerations

Important guidance for competition policy is offered by the economic theory of the "second best". Laffont puts forward the basic argument on the second best in the following terms:

> Competition is an unambiguously good thing in the first-best world of economists. That world assumes large numbers of participants in all markets, no public goods, no externalities, no information asymmetries, no natural monopolies, complete mar-

> kets, fully rational economic agents, a benevolent court system to enforce contracts, and a benevolent government providing lump sum transfers to achieve any desirable redistribution. Because developing countries are so far from this ideal world, it is not always the case that competition should be encouraged in these countries. (Laffont, 1998, p. 237)

The basic idea is, of course, that, if some of the conditions for a competitive equilibrium are violated, a second-best solution would involve restrictions on competition. Precisely what those restrictions should be is a much more difficult question, because that depends on the nature and structure of the existing distortions and whether these can be remedied through other means. Laffont is quite pessimistic about developing countries being able to implement competition laws because of widespread rent-seeking, corruption and ineffective governments in these countries. He makes a valid point that implementation of competition law requires a strong state which many developing countries lack.

Pessimism is not, however, warranted in the case of all developing countries. Many semi-industrial countries have strong and effective governments, though not always fully democratic. These include some of the most populous countries in the developing world, such as China, India, Brazil and Mexico. There are also several well-known developmental states. Moreover, the question of corruption should be kept in perspective. Many of the East Asian countries, including China, do not rank very high in the transparency league table, and yet, these are the countries with the fastest rates of sustained growth in the history of humankind. Indonesia, for example, has become a byword for corruption during Suharto's regime, yet during his thirty-year rule the country's record was the best in the world in reducing poverty (Stiglitz, 1998).[25] In South Korea, the two presidents who presided over that country's rapid industrial expansion in the late 1970s and 1980s were each convicted by courts for accepting hundreds of millions of US dollars in bribes. Even in relation to developed countries, the post-WWII Italian economic miracle does not seem to have been hampered by widespread cronyism and corruption within the political establishment. There is clearly no simple relationship between corruption, economic growth and a country's ability to implement interventionist economic and industrial policy, including competition policy.

What kind of competition policy would be appropriate for those developing countries with reasonably effective states as well as the necessary institutional framework to carry out a developmental program? The central point here is that the second-best framework outlined above is much too narrow for taking into account the developmental dimension. This is in part because for a developing country the purpose of competition policy cannot simply be the promotion of competition as a good thing *per se*, but to foster economic development. This would in some instances involve restriction of competition and in others its

vigourous promotion. In order to raise the living standards of their people over time, developing countries need high rates of investment to achieve fast growth of productivity. High rates of investment in turn normally require reasonable, if not high, rates of profits in order to maintain the private sector's propensity to invest. This consideration leads to the view that there may at times be too much competition rather than too little. Competition would be too much if it leads to price wars and sharp falls in profits, which are likely to diminish the corporate desire to invest. In the real world of incomplete and missing markets, which is particularly the case in developing countries, the latter may also require government co-ordination of investment decisions to prevent over-capacity. A developing country cannot therefore afford to have maximum competition; rather it must operate with an optimal degree of competition or with an appropriate blend of competition and co-operation to achieve its long term goals of faster and sustained economic growth. As was noted earlier in the discussion in Section III, this is also the conclusion that emerges from new developments in the theory of industrial organisation.

To sum up, the above analysis suggests that competition policy cannot be a unique, one-size-fits-all policy that is appropriate for all developing countries. The optimal policy will differ between countries depending on their stage of development and the effectiveness of their governments as well as their supporting institutional frameworks.

4.b. Competition Policies in Advanced Countries

It may be useful to consider briefly what the nature and practice of competition policy has been in developed countries. What lessons can developing countries draw from their experience?

4.b.1. United States

The United States is the country with the longest history of antitrust laws and laws prohibiting restraints on competition. In the period from the end of WWII to the 1980s, with respect to the former, the US followed a structural policy which more or less forbade mergers in the same industry. This is thought by some to have encouraged the conglomerate mergers of the 1960s. With the liberalisation of the world economy and the US's difficulties in maintaining equilibrium in its current account, there appears to have been a relaxation of the strict interpretation of the competition laws. It is a moot point whether this was due more to the influence of foreign competition or to that of the Chicago School, but the upshot was that the enforcement of competition laws became relatively relaxed. For example, the Federal Trade Commission began to take account of economies of scale as a defence against charges of increased market power. More recently, there have been further relaxations of antitrust laws in

the light of advances in economic theory and the courts' acceptance of these. In a recent review, Jonathan Baker (1999) concluded as follows:

> Three decades ago, antitrust law relied heavily upon "per se" rules, which took the broad brush approach of deeming certain classes of business practices anticompetitive without regard to their effects in any particular case. Today, a case-by-case analysis is more common, often under the judicial rubric of applying the "rule of reason". (Baker, 1999, p. 191)

The *per se* rules, which prevailed for a long time in the US conception and implementation of antitrust law, reflected the belief that competition is a good thing *per se*, without regard to its economic consequences. This is the doctrine that is now changing.

4.b.2. European Union

The European Union's competition law consists of Article 81 and 82 of the Treaty of Rome and the national competition laws of the member states. The primary objective of these laws is the creation of the single European market. However, European competition law also makes provision for industrial policy under strict guidelines, as well as provision for other objectives such as fairness, equality of income distribution and other social goals (e.g. reducing regional disparities and unemployment).

Audretsch, Baumol and Burke (2001) note two shortcomings in EU competition law from the perspective of dynamic efficiency: the lack of clarity on the social welfare objective of the laws and an emphasis on static efficiency. They argue that the economic prescription for competition policy is relatively simple only if one ignores such phenomena as: (a) the variation in the abilities of firms to exploit particular profit opportunities; (b) the evolution of such capability with the passage of time; and (c) the manipulation of barriers to entry or the incentives for innovation and its possible abuse as a means to undermine competition. (Audretsch *et al.*, 2001, p. 629).

4.b.3. Japan

Competition policy in Japan has evolved over time since its inception under the US military occupation in the late 1940s. The period which is relevant for developing countries is that from 1950–1973, when Japan was much more like a newly-industrialising country than it is today. During this period Japan achieved extraordinarily fast economic growth, with manufacturing production rising at over 13% per annum, GDP at 10% a year and its share of world exports rising by a huge ten percentage points.

At this time, Japanese industrial policy, formulated by the Ministry of International Trade and Industry (MITI), had very much the upper hand over the Fair Trade Commission (FTC), the competition watchdog. One of MITI's main ob-

Table 12.12 Japanese cartel agreements exempted from the anti-monopoly law by the Fair Trade Commission or competent Ministry by exempting statute: 1964–1973[a]

Statutory basis for exemption	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973
Depression cartels	2	2	16	1	0	0	0	0	9	2
Rationalisation cartels	14	14	14	13	13	12	10	13	10	10
Export cartels	201	208	211	206	213	217	214	192	175	180
Import cartels	1	2	3	4	3	4	4	3	2	2
Cartels under Medium and Small Enterprises Organisation Act	588	587	652	634	582	522	469	439	604	607
Cartels under Environment Sanitation Act	106	122	123	123	123	123	123	123	123	123
Cartels under Coastal Shipping Association Act	15	14	16	15	22	22	22	21	19	19
Cartels under other statues	43	50	44	44	47	48	56	53	34	42
Total	970	999	1079	1040	1003	948	898	844	976	985

Note: [a] Number in force in March of each year.

Source: Japanese Fair Trade Commission (1973), p. 27.

jectives was to ensure a high rate of profitability and investment in Japanese industry. MITI was therefore always concerned with questions of "ruinous competition" leading to reduced profits and a lower propensity to invest. The Ministry thus officially sponsored a wide variety of cartels (including recession cartels, export cartels and technology cartels, to name a few), sequenced investment by firms and intervened in the exit and entry decisions of firms, all of which contributed to the high concentration ratios observed in the Japanese economy (see tables 12.1 and 12.12).

Some scholars, such as Caves and Uekusa (1976), have been stringent in their criticism of this weak competition policy arguing, that it has imposed serious allocational inefficiencies on the Japanese economy.

However, MITI did not just thwart the FTC's codes and objectives, but it also implemented an industrial policy that encouraged contest-based competition between oligopolistic firms where the rewards were access to cheap credit and foreign exchange as well as, where necessary, protection from international competition. These rewards were contingent on relative performance either in export markets, technological development, or in introducing new products. The result was, as Odagiri (1994) and Porter (1990) note, that rivalry between firms in Japan was extremely intense. Indeed, as the persistency of profitability studies of the kind reviewed in Section II indicate, the intensity of competition in Japan's manufacturing sector has been greater than in US manufacturing (Odagiri, 1994).

Thus, Japan followed a policy that promoted dynamic efficiency (in the sense of maximising long-term growth of productivity) through an institutional structure that combined both co-operation and competition between firms. This policy has much to commend it to developing countries. It is fully in accord with the analytical considerations for an appropriate competition policy for developing countries outlined earlier and is also consistent with the latest advances in economic theory.[26]

4.c. New Concepts for Competition Policy for Economic Development

In Singh and Dhumale (1999), we expressed serious misgivings about the WTO Working Group's analysis of competition policy for developing countries. It did not seem to us to meet one of the Group's chief objectives: to take the development dimension of competition policy fully into account. We came to the view that a discussion on competition policy and economic development in terms of the WTO concepts such as market access, reciprocity and national treatment was prejudicial to the interests of developing countries. To take the development dimension properly into account, it was essential to have new definitions and fresh concepts rather than to conduct the exercise in terms of the existing WTO terminology.

On the basis of the modern theory of industrial organisation, as well as the history of competition policy in developed countries, Singh and Dhumale suggested that development-friendly competition policies need to have different objectives from those normally posited for advanced economies. Further, such policies also need to be specific to the stage of a country's economic and industrial development as well as its institutional and governance capacities. In relation to the WTO Working Group's tasks, this analysis suggested the following concepts to address the developmental dimensions of competition policy:

- the need to emphasise dynamic rather than static efficiency as the main purpose of competition policy from the perspective of economic development;
- the concept of "optimal degree of competition" (as opposed to maximum or perfect competition) to promote long term growth of productivity;
- the related concept of "optimal combination of competition and co-operation" to achieve fast long term economic growth;
- the critical significance of maintaining the private sector's propensity to invest at high levels, and hence the need for a steady growth of profits; the latter in turn may necessitate government co-ordination of investment decisions so as to prevent over-capacity and falling profits;
- the concept of simulated competition, i.e., contests, for state support, which can be as powerful as real market competition;
- the crucial importance of industrial policy in achieving the structural changes required for economic development; this in turn requires coherence between industrial and competition policies.

The development dimension is thus far from being fully taken into account by suggestions that all that developing countries need is a longer time frame to be able to implement the US or UK type of competition policy. The special and different circumstances of developing countries and their developmental needs require a creative application of the concepts above to competition policy questions.

5. MULTILATERAL COMPETITION POLICY VERSUS INTERNATIONAL COMPETITION AUTHORITY

At the WTO a number of advanced countries have been pressing developing countries to negotiate to make competition policy subject to that organisation's multilateral disciplines, so as to ensure "fair play" and "level playing fields" between countries.[27]

Developing countries have been opposed to such proposals. Their formal stance has been to suggest that as many of them have no experience of competition policy, they are not in a position to be able to enter into negotiations on these matters. The real reason for developing countries' opposition is that they do not wish any new disciplines to be included in the WTO agreements because of the provision of cross-sanctions: a violation in one area may be penalised in another by the complaining country (if the complaint is held to be justified). Until the Doha meeting developing countries took the view that the Uruguay Round Agreements (that established the WTO) needed to be properly reviewed for their impact on economic development before undertaking a new round of tariff cutting or starting negotiations on new subjects such as competition policy and the multilateral agreement on investment. However, after the Doha Ministerial meeting developing countries may find it difficult to maintain such a stance for long.

It may be interesting to observe that there has been an ironic reversal of roles here. In the past, developing countries were in favour of multilateral action to restrict business practices of the large multinational companies. At the insistence of developing countries the UN General Assembly in December 1980 adopted by Resolution 35/63 a "Set of Multilaterally Agreed Equitable Principles and Rules for the Control of Restrictive Business Practices". The "Set" is fairly comprehensive in scope and covers a wide range of restrictive business practices by multinationals, including the abuse of their dominant positions, whether achieved through mergers and acquisitions or joint ventures. At that time, developing countries were in favour of making SET legally binding. This, however, was not acceptable to developed countries. Today, the position is the other way around, with advanced countries seeking a binding multilateral agreement through the WTO and developing countries opposing it.

Proponents of a multilateral agreement on competition policy have put forward the following arguments in its favour:

i) It would be helpful to developing countries as it would enable them to restrain anti-competitive behaviour and cartelisation by large, advanced country corporations.
ii) It may help to bring the TRIPS agreement under multilateral competition disciplines. Maskus and Lahouel (2000) suggest that the possible abuse of intellectual property rights, as well as parallel imports, could be regulated by a multilaterally agreed competition policy.
iii) Stiglitz (1999) suggests that if there were a new multilateral competition policy agreement this would help to blunt the potency of anti-dumping laws by bringing them into the normal framework of predation under competition laws. The predation test is much stricter than the anti-dumping measures that countries have been using under the WT).[29]

iv) A multilateral competition policy will help foster competition both nationally and internationally, from which it is suggested that developing countries would greatly benefit. Perroni and Whalley (1998) quantify the potential gains of developing countries from the introduction of disciplines on competition, suggesting that the potential gains for developing countries could be large, perhaps in the region of 5–6% of national income. This would make a competition policy negotiation of potentially more significance to developing countries than the whole of the trade disciplines achieved in the Uruguay Round (Perroni and Whalley, 1998, p. 493).[30]

These gains would include those stemming from the replacement of anti-dumping measures by competition law, reduction of mark-ups of foreign suppliers and reduced concentration in domestic markets.

There are, however, powerful arguments against multilateral disciplines from the perspective of developing countries. The first is that a multilateral agreement on competition policy, to be development-friendly, must be highly flexible, allowing each country to determine its competition policy for itself on the basis of the country's needs and circumstances. This implies that if the cost-benefit analysis for a particular country shows there is no gain from it, the country need not have a competition policy at all.

Critics of a multilateral competition policy also suggest that a main motivation for developed countries to seek a competition policy agreement is for market access to developing countries. Developed countries would like to have, in addition to an agreement on competition policy, an international agreement on foreign direct investment (FDI). Under the latter, large advanced country multinationals would be permitted to invest anywhere they like in any quantity and at any time without any let or hindrance from developing country governments. In addition, once established, the multinationals would have "national treatment", i.e. be treated the same as national firms. An ambitious multilateral agreement on these issues would accord multinationals equal treatment in both pre and post-establishment phases.

However, such an agreement would be seriously prejudicial to economic development. In a detailed analysis of FDI as a source of long-term finance for developing countries, Singh (2001) has argued that unless it is adequately regulated by their governments, in the particular circumstances of these countries, where they are subject to frequent internal and external shocks, it would lead to short- and long-term financial fragility. To avoid this fragility, it is necessary for developing country governments to control: (a) the timing of the FDI; (b) the total amount of FDI; as well as (c) the selection of large projects by multinationals. These measures are needed to ensure that there is no mismatch of the time profile of a country's foreign exchange inflows and outflows. Such time inconsistency can lead to a liquidity crisis, which, as the experience of Asian

economic crisis shows, may degenerate into solvency problems with ultimately devastating consequences for the real economy.[31]

Multinationals often complain that there is no "level playing field" between them and the national corporations which are government supported; hence, the multinationals' demand for "national treatment". However, the actual situation is quite the opposite: the playing fields are tilted heavily in favour of multinationals who have considerable market power in markets for outputs as well as inputs. The current international merger movement is making these fields even more unequal even from the perspective of the *large* developing country corporations.

The mechanical application of the WTO principle of "national treatment" in these circumstances would clearly lead to perverse results that would both harm economic development in developing countries as well as lead to global economic inefficiency. The magnitude of the latter would be determined by the extent to which the multinationals' financial advantage over domestic firms arises from market power rather than from genuine economies of scale.

To provide a simple illustration, it should be perfectly legitimate for a developing country competition authority to allow large domestic firms to merge so that they can go some way toward competing on more equal terms with multinationals from abroad. Even if the amalgamating national firms are on the horizontal part of the L-shaped static cost curve, bigger size may still promote dynamic efficiency for the reason that firms need to achieve a minimum threshold size to finance their own R&D activities. The competition authority may therefore quite reasonably deny national treatment to the multinationals and prohibit their merger activity (because they are already large enough to achieve either static or dynamic economies of scale in this sense). In these circumstances, a violation of the doctrine of national treatment is likely to be beneficial both to economic development and to competition.

In view of these serious limitations of multilateral competition policy it is essential to look for alternative means of international co-operation on this subject. This is because, as argued earlier, even if developing countries had development friendly national competition policies, they would still need international assistance to restrain the anti-competitive conduct of dominant multinationals as well as to limit the adverse effects of mega-mergers associated with the merger movement of the 1990s. The best way, it seems to me, to provide such help would be through an International Competition Authority. The characteristics and responsibilities of this Authority would include the following:

- It would be charged with maintaining fair competition in the world economy and keeping the markets contestable by ensuring that the barriers-to-entry to late industrialisers are kept at low levels.

- Analogous to the social welfare objectives of the European Commission, the proposed International Authority would be asked to pay attention to the special needs of the developing countries, to competitive opportunities for small and medium sized firms, to facilitate transfer of technology to developing countries and to ensure fair prices and fair distribution of wealth.
- It would have the authority to scrutinise mega-mergers and to deter the mega-firms from abusing their dominant position.
- Again on the European Union model, the International Competition Authority would be concerned mainly with cross-border or international aspects of the workings of competition. Below the authority, at a national level, the member countries would have their own national competition policies.
- For good administrative and practical reasons, references to the competition authority would only be permissible in case of anti-competitive behaviour by corporations above a certain size. The size criterion would normally keep even most large developing country corporations outside the direct purview of the competition authority.
- In relation to the international merger movement, the authority would attempt to limit growth by merger by large multinationals under its purview. They would be allowed to merge provided they divested themselves of a subsidiary of equal value. This would mean that multinationals would not be able to grow by mergers, but they could expand through organic growth or green-field investment. It would not stop them from taking over other firms, provided they were willing to sever a similar sized subsidiary.
- In the light of the extended discussion of the international merger movement in Section III, the main merits of this proposal are as follows. As mergers, on average, do not appear to improve economic efficiency, and the mega-mergers have the potential of increasing market dominance and reducing contestability, discouraging such mergers would therefore enhance global competition and global economic efficiency while at the same time being distibutionally more equitable.
- The governance of the ICA would have proper representation of developing countries and would not be dominated by developed countries.

Although international co-operation on competition policy, in the form outlined above, would be of particular benefit to developing countries, it also has useful features to assist the large multinational corporations. The International Competition Authority would, for example, be able to provide multinationals under its purview with unambiguous decisions on mergers and other competition related matters. Instead of being subject to the often conflicting decisions of

many different jurisdictions (e.g., the US, the European Union, Japan, and overtime countries like India and China) International Competition Authority's rulings would prevail overall national and regional jurisdiction.[32]

There is no illusion that an international agreement of the above kind would immediately be acceptable to advanced countries. Nevertheless, it indicates the nature of economic arrangements in this area that would best serve developmental needs of poor countries. It may, however, be helpful to proceed to the establishment of the ICA in stages. At the first stage, the authority may have no coercive powers, but simply be able to monitor and to report on abuses of dominant market positions, on mergers, and the authority's other competition objectives.[33] Such monitoring would itself be beneficial to developing countries, as it would provide them with information on cartels and on market power abuses of multinationals. Developing countries would find it difficult to acquire such information otherwise. With the experience gained from this kind of limited international co-operation, nations can, over time, work towards greater co-operation by giving ICA the necessary powers to enforce its rules.

There is finally the question whether ICA should nevertheless be an integral part of the WTO or if it should be a stand-alone authority. In addition to the reasons mentioned earlier in the discussion of the multilateral competition policy, there are also other considerations that would suggest the latter would be the better option. This is in part because questions of competition policy go much beyond those related to international trade. Further, WTO does not have the expertise to be the world's "FTC". Moreover, the primary objectives of competition policy tend to be rather different from those of the promotion of free trade through measures such as market access and national treatment. Since, as indicated above, the latter concepts are not very helpful to developing countries, it would be best to keep the two institutions (the WTO and the ICA) separate.

6. CONCLUSION

The main points of this chapter have been presented in the Introduction. Its central message is to suggest that developing countries at the WTO are faced with a serious difficulty in discussions on competition policy as well as on other similar issues as long as the whole discourse is expressed in terms of the WTO concepts and language. These are inadequate to reflect the developmental concerns of emerging countries. Developing countries need to develop the appropriate language and concepts within which their concerns can be properly articulated. Hopefully this chapter has made a small contribution in that direction.

The Preamble to the WTO notes that "trade and economic endeavour should be conducted with a view to raising the standards of living, ensuring full em-

ployment and a large and steadily growing volume of real income and effective demand". It is further stated that "there is need for positive efforts designed to ensure that developing countries, and especially the least-developed among them, secure a share in the growth in international trade commensurate with the needs of their economic development" (quoted in Rodrik, 2001). Full employment and economic development are not only the ultimate goals of the WTO but these have also been repeatedly endorsed by the international community. In 1995, 117 Heads of State or Government attending the Copenhagen Social Summit endorsed the Copenhagen Declaration, which put primary emphasis on the promotion of full employment and poverty reduction. More recently, similar declarations have been made at the Millennium Summit at the UN and other fora. Indeed, the right to a decent living has virtually acquired the status of a universal human right.

If experience and analysis show that the primary goals of the WTO are being harmed rather than helped by specific measures such as TRIMS, or the equal application to all countries of a particular procedural principle such as national treatment, it is the latter which should be changed. It is the primary goals rather than the procedural rules of an international organisation that should dominate, especially as the former are widely endorsed by the world community as a whole.

In this spirit the chapter has put forward a proposal for a development friendly International Competition Authority in order to control anti-competitive conduct of the world's large multinational corporations (above a certain threshold of size), as well as to control their propensity to grow by take-overs and mergers. In order to maintain contestability and efficiency of international markets, it is proposed that the large multinationals should be allowed to take over another company only if they sell off a subsidiary of similar value. Thus, even the largest multinationals are not stopped from growing provided they expand their size by green-field investment. Neither are they stopped from taking over other firms, provided they are able to sell off equal value subsidiaries, i.e. they cannot grow by mergers or take-overs. It is argued here that these institutional arrangements would both be more efficient as well as more equitable compared with the present situation. It is, however, recognised that the advanced countries are not yet ready to cede some of their sovereignty that would be required for such close international co-operation. The evolution towards the establishment of the ICA could, therefore, occur in stages. As a first step the Authority could be entrusted only with fact-finding and monitoring anti-competitive behaviour and threats to the contestability of international markets. This could evolve over time into deeper North-South co-operation and the full fledged establishment of the ICA according to the principles outlined in the chapter.

NOTES

* This chapter was initially published as Harvard University Center for International Development, United Nations Conference on Trade and Development, Research paper for the Intergovernmental Group of Twenty-Four on International Monetary Affairs, No. 18, September 2002. The author is grateful to Professor Dani Rodrik and the head of UNCTAD, Harvard University, for permission to use this material for this chapter.

† Tony Thirlwall's and my research have overlapped extensively throughout our careers. We have both worked on de-industrialisation and structural changes in developed and developing countries and have both been disciples of Kaldor. It is therefore a pleasure to dedicate this article to Tony Thirlwall to mark his retirement. In the international debate on economic development, competition and competition policies are new issues for developing countries but have enormous implications for industrialisation and economic development.

1. As is explained later, the WTO Ministerial Meeting at Cancun in 2003 broke down over the Singapore issues and subsequently these were temporarily withdrawn by the North from the negotiation process. However, this is a tactical retreat by the North which may re-open the questions whenever they think it opportune.
2. This WTO Working Group on Trade and Competition Policy (WGTCP) has been meeting under the chairmanship of Professor Frédéric Jenny for the last four years. Their work will be commented upon later in the chapter.
3. The US takes a more measured stance. The US Trade Representative's Fact Sheet summarising the results of the Doha ministerial noted in relation to competition policy that a two-stage "modest" negotiation was agreed upon. The first stage would seek clarification of "core principles", including transparency, non-discrimination and procedural fairness. At the second stage the "timing and specific content" of the negotiations will be decided.
4. The account of the Doha ministerial meeting in this paragraph and the following one comes from *Bridges* (International Centre for Trade and Sustainable Development), Year 5, No. 9, November/December 2001, p. 6.
5. Apart from the WTO, the CUTS Centre for International Trade, Economics and Environment reports that competition policy is the on the agenda of the proposed Free Trade Area of the Americas (FTAA) as well as the European Union / Africa, Caribbean, Pacific Grouping (EU / ACP).
6. For the structural analysis of the Asian crisis, see for example Greenspan (1998), Phelps (1999), and IMF (1998). The structural analysis of the Asian crisis is of course neither necessarily accurate nor universally accepted. For a strong rebuttal see, Singh (1999) and Singh and Weisse (1999); for alternative analyses, see among others Radelet and Sachs (1998); Feldstein (1998, 2002); Stiglitz (1999).
7. See de Soto (1989).
8. The first bias arises from the fact that the data in Table 12.2 pertain to size of plants rather than enterprises. This bias would, however, understate relatively the small firms' contribution to the economies of developing countries. This is because there are likely to be more multi-plant enterprises in developed than in developing countries. Secondly, the data by a large consider only the formal sector. The informal sector in developing countries is typically considerably larger than that in advanced countries. This bias would also therefore operate in the same direction as the first one.
9. A distinction is often made between competition law and competition policy – the latter being a wider concept encompassing elements of industrial policy among other things – see for example Hoekman and Kostecki (2001). Here, formal competition policy is used in the narrow sense. The broader concept employed here is that of industrial policy.
10. These Estimates are reported in World Bank (2002, p. 141).
11. For a fuller discussion of these issues, see World Bank (1993), Amsden (2001). For a theoretical analysis, see Stiglitz and Nalebuff (1983).
12. This is a rather different conception of competition from that involved in the Arrow-Debreu formulation of general equilibrium in a decentralised market economy. In this conception, a competitive equilibrium exists under certain specified conditions that leads to a Pareto-optimal

allocation of resources. However, the notion of competition described in the text is rather different: it is concerned with the common-sense understanding of competition as an incentive to elicit maximum individual or organisation effort.

13. See further Amsden and Singh (1994), Laffont (1998).
14. See further Vickers (1994), and Nickell (1996).
15. There is a vast literature on this subject. See Amsden (1989), Rodrik (1994), Singh (1995), and Wade (1990) among others.
16. The classic reference here is Piore and Sabel (1984). See also Best (1990).
17. See Pollit (1999) for a useful review of the literature on the subject.
18. The UK had a similarly large merger wave roughly around the same time. Parenthetically, the best historical data on mergers is available for the US and the UK. The two countries have similar institutions and corporate law, which allows useful comparisons between them. International comparisons involving other countries are always more difficult because of differences in corporate law as well as the definition of what constitutes a "merger". See further Hughes and Singh (1980).
19. Best (1990), p. 104.
20. For recent reviews of the literature see Mueller (1997) and Tichy (2002).
21. A few studies have found a small increase in profitability after mergers. However, 90% of studies have found either no change or reduced profits.
22. For a general discussion and quantitative analysis of the cartels that have been detected and presented during the last two decades, see Evenett, Levenstein and Suslow (2001).
23. On the incidence of cartels in the 1930s and the post-war period, see Mason (1946) and Scherer (1994). On the reasons for the rise in mergers and take-overs in the post-war period, see Singh (1992, 1993).
24. See further *Economist* (2002).
25. As Stiglitz rightly points out, one needs to distinguish between a Suharto and a Mobuto.
26. According to Audretsch, Baumol and Burke (2001), the dynamic efficiency perspective of competition policy and modern theoretical analysis requires consideration of issues such as the appropriate ease of entry; appropriate inter-firm coordination; innovation, trade and monopoly power; anti-competitive innovation; monopoly in innovative markets; and price discrimination when R&D costs are substantial and continuing. In other words, a host of dynamic factors must be analysed. This is the authors' recommendation for an appropriate competition policy in advanced countries. What MITI did in Japan during its period of high growth was to adapt such considerations to their particular circumstances.
27. Subsequently the Singapore issues became the subject of contention between a much more organised South and the North. This led to the breakdown of the Ministerial meeting at Cancun in 2003 and also led to the subsequent withdrawal of the Singapore issues from the immediate ambit of negotiations between the North and the South. Nevertheless, the question of competition policy remains important in its own right, whether or not it is regarded as a part of the WTO agenda. This issue is discussed in the text further on.
28. Similarly in 1986, the OECD issued guidelines concerning restrictive business practices by multinationals. Under these guidelines, which again were advisory rather than legally enforceable, multinational enterprises were enjoined to refrain from a wide range of anti-competitive activities including abuses of intellectual property rights, predatory behaviour, competition reducing acquisitions, etc. (see further, OECD 1986; Scherer, 1994).
29. Stiglitz presents a recent report on US anti-dumping cases that suggests that, if these had been subject to the equivalent US competition policy standard of predation, more than 90% of them would have failed. (Stiglitz, 1999).
30. Quoted in Correa (1999).
31. These arguments for permitting developing countries to be able to monitor and to regulate FDI flows are further complemented by considerations of technology transfer and spillover benefits. Research shows that both of those occur best when FDI is carefully regulated and fits well into a country's development program. See further Singh (2001), Singh (2003).
32. For the difficulties involved for corporate decision making as a consequence of overlapping jurisdictions of the competition authorities of different countries, See Trade and Competition: Towards a global response. http://trade-info.cec.eu.int/europa/2001newround/com.htm.

33. Scherer, 1994, makes a similar point in relation to his proposal for an international agreement on competition policy.

REFERENCES

Amsden, A. (1989), *Asia's Next Giant: South Korea and Late Industrialization*, New York: Oxford University Press.

Amsden, A. (2001), *The Rise of "The Rest": Challenges to the West from Late-Industrializing Economies*, New York: Oxford University Press.

Amsden, A. and A. Singh (1994), "The Optimal Degree of Competition and Dynamic Efficiency in Japan and Korea", *European Economic Review*, **28**, 941–51.

Audretsch, D., W. Baumol and A. Burke (2001), "Competition Policy in Dynamic Markets", *International Journal of Industrial Organization*, **19**(5) , April, 613–34.

Aw, B., S. Chung and M. Roberts (2002), "Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufacturers", *NBER Working Paper*, 8766, February.

Baker, J. (1999), "Developments in Antitrust Economics", *Journal of Economic Perspectives*, **13**(1), Winter, 181–94.

Baumol, W. (2001), "When is Inter-firm Coordination Beneficial? The Case of Innovation", *International Journal of Industrial Organization*, **19**, 727–37.

Best, M. (1990), *The New Competition: Institutions of Industrial Restructuring*, Cambridge, MA: Harvard University Press.

Caves, R. and M. Uekusa (1976), *Industrial Organisation in Japan*, Washington, DC: Brookings Institution.

Correa, Carlos M. (1999), "Competition Law and Development Policies", paper presented at the Zurich Seminar on Issues of International Competition Law, University of Zurich, 8–10 July.

CUTS (2000), "Globalisation, Competition Policy and International Trade Negotiations: Considerations for Developing Countries", monographs on Investment and Competition Policy, No. 4, Centre for International Trade, Economics and Environment.

de Soto, H. (1989), *The Other Path*: *The Invisible Revolution in the Third World*, New York: Harper & Row.

Economist, The (2002), "Cartels: Fixing for a Fight", 20 April, pp. 71–2.

Evenett, Simon J., Margaret C. Levenstein and Valerie Y. Suslow (2001), "International Cartel Enforcement: Lessons from the 1990s", *World Economy*, **24**(9), 1221–45.

Feldstein, M. (1998), "Trying to Do Too Much", *Financial Times*, 3 March.

Feldstein, M. (2002), "Economic and Financial Crises in Emerging Market Economies: Overview of Prevention and Management", *NBER Working Paper*, 8837, March.

Glen, J., K. Lee and A. Singh (2001), "Persistence of Profitability and Competition in Emerging Markets", *Economics Letters*, **72**, 247–53.

Glen, J., K. Lee and A. Singh (2002), "Corporate Profitability and the Dynamics of Competition in Emerging Markets: A Time Series Analysis", *The Economic Journal*, **113**(491), November, F465–F484.

Goddard, J. and J. Wilson (1999), "The Persistence of Profit: A New Empirical Interpretation", *International Journal of Industrial Organization*, **17**, 663–87.

Gort, M. (1969), "An Economic Disturbance Theory of Mergers", *Quarterly Journal of Economics*, **83**, 624–42.

Greenspan, A. (1998), *Testimony before the Committee on Banking and Financial Services*, United States House of Representatives, 30 January.

Gugler, K., D. Mueller, B. Burcin Yurtoglu and C. Zulehner (2003), "The Effects of Mergers: An International Comparison", *International Journal of Industrial Organisation*, **21**, 625–53.
Hoekman, B.M. and M.M. Kostecki (2001), *The Political Economy of the World Trading System*, Oxford: Oxford University Press.
Holmstrom, B. and S. Kaplan (2001), "Corporate Governance and Merger Activity in the United States: Making Sense of the 1980s and 1990s", *Journal of Economic Perspectives*, **15**(2), Spring, 121–44.
Hughes, A. and A. Singh (1980), "Mergers, Concentration and Competition in Advanced Capitalist Economies: An International Perspective", in D. Mueller (ed.), *The Determinants and Effects of Mergers*, Cambridge, MA: Oelgeschlager, Gunn and Hain, pp. 1–26.
I.M.F. (1998), "The Asian Crisis: Causes and Cures", *Finance and Development*, **35**(2), June.
International Centre for Trade and Sustainable Development (ICTSD) (2001), *Bridges Weekly*, **5**(39), 15 November.
Japanese Fair Trade Commission (1973), *The Antimonopoly Act of Japan*, Tokyo: Staff Office.
Laffont, J.-J. (1998), "Competition, Information, and Development", *Annual World Bank Conference on Development Economics*, *1998*, 237–57.
Mason, Edward S. (1946), *Controlling World Trade: Cartels and Commodity Agreements*, New York: McGraw-Hill.
Maskus, K. and M. Lahouel (2000), "Competition Policy and Intellectual Property Rights in Developing Countries", *World Economy*, 595–611.
Mueller, D. (1997), "Merger Policy in the United States: A Reconsideration", *Review of Industrial Organization*, **12**(5–6), 655–85.
Nalebuff, B. and J. Stiglitz (1983a), "Information, Competition, and Markets", *The American Economic Review*, **73**(2), 278–83.
Nalebuff, B. and J. Stiglitz (1983b), "Prizes and incentives: towards a general theory of compensation and competition", *Bell Journal of Economics*, **14**, Spring, 21–43.
Nickell, S. (1996), "Competition and Corporate Performance", *Journal of Political Economy*, **104**(4), 724–46.
Odagiri, H. (1994), *Growth Through Competition, Competition Through Growth: Strategic Management and the Economy in Japan*, New York: Oxford University Press.
OECD (1986), *The OECD Guidelines for Multinational Enterprises*, Paris: OECD.
Perroni, C. and J. Whalley (1998), "Competition Policy and the Developing Countries", in H. Thomas and J. Whalley (eds), *Uruguay Round Results and the Emerging Trade Agenda: Quantitative-based Analyses from the Development Perspective*, New York and Geneva: UNCTAD.
Phelps, E.S. (1999), "The Global Crisis of Corporatism", *Wall Street Journal*, 25 March.
Piore, M. and C. Sabel (1984), *The Second Industrial Divide*, New York: Basic Books.
Pollitt, M. (1999), "A Survey of the Liberalisation of Public Enterprises in the UK since 1979", *DAE Working Paper*, no. 9901, Department of Applied Economics, University of Cambridge.
Porter, M. (1990), *The Competitive Advantage of Nations*, London; Macmillan.
Radelet, W. and J. Sachs (1998), "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects", *Brookings Papers on Economic Activity*.
Rodrik, D. (1994), "King Kong Meets Godzilla: The World Bank and The East Asian

Miracle", in A. Fishlow (ed.), *Miracle or Design? Lessons from the East Asian Experience*, Washington, DC: Overseas Development Council, Policy Essay No. 11.
Rodrik, D. (2001), "The Global Governance of Trade: As if Development Really Mattered", *UNDP Background Paper*, New York: UNDP.
Scherer, F.M. (1994), *Competition Policies for an Integrated World Economy*, Washington, DC: Brookings Institution.
Shleifer, A. and L.H. Summers (1988), "Breach of Trust in Hostile Takeovers" in Auerbach, A.J. (ed.), *Corporate Takeovers: Causes and Consequences*, Chicago: National Bureau of Economic Research, University of Chicago Press.
Singh, A. (1971), *Take-overs: Their Relevance to the Stock Market and the Theory of the Firm*, Cambridge: Cambridge University Press.
Singh, A. (1975), "Take-overs, Economic Natural Selection and the Theory of the Firm", *Economic Journal*, September.
Singh, A. (1992), "Corporate Takeovers", in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave Dictionary of Money and Finance*, London: Macmillan, pp. 480–86.
Singh, A. (1993), "Regulation of Mergers: A New Agenda", in R. Sugden (ed.), *Industrial Economic Regulation: A Framework and Exploration*, London: Routledge, pp. 141–58.
Singh, A. (1995), *Corporate Financial Patterns in Industrialising Economies: A Comparative International Study*, Technical Paper 2, International Finance Corporation, Washington, DC: World Bank.
Singh, A. (1998), "Liberalization, the Stock Market and the Market for Corporate Control: A Bridge Too Far for the Indian Economy?", in I.J. Ahluwalia and I.M.D. Little (eds), *India's Economic Reforms and Development: Essays for Manmohan Singh*, New Delhi: Oxford University Press.
Singh, A. (1999), "'Asian Capitalism' and the Financial Crisis", in J. Michie and J. Grieve-Smith (eds), *Global Instability*, London: Routledge.
Singh, A. (2001a), "Foreign Direct Investment and International Agreements: A South Perspective", *T.R.A.D.E. Occasional Paper*, no. 6, October, Geneva: South Centre.
Singh A. (2001b), "Statement made to the German Parliament's Study Commission on 'Globalisation of the World Economy – Challenges and Responses", 14 May.
Singh, A. (2003a), "Capital Account Liberalization, Free Long-Term Capital Flows, Financial Crises and Economic Development", *Eastern Economic Journal*, **29**(2), 191–216.
Singh, A. (2003b), "Competition, Corporate Governance and Selection in Emerging Markets", *The Economic Journal*, **113**(491), November, F443–F464.
Singh, A. and R. Dhumale (1999), "Competition Policy, Development and Developing Countries", *T.R.A.D.E. Working Papers*, November, Geneva: South Centre, (available at www.southcentre.org/publications/index.htm). Subsequently published in P. Arestis, P.M. Baddeley and J. McCombie (eds) (2001), *What Global Economic Crisis?*, Palgrave, London, pp. 122–45.
Singh, A. and Weisse, B. (1999), "The Asian Model: A Crisis Foretold?", *International Social Science Journal*, **160**, 203–15.
Stigler, G.J. (1950), "Monopoly and Oligopoly by Merger", *American Economic Review*, Supplement, May.
Stiglitz, J. (1998), "Restoring the Asian Miracle", *Wall Street Journal Europe*, 3 February, p. 4.
Stiglitz, J. (1999), "Reforming the Global Financial Architecture: Lessons from Recent Crises", *Journal of Finance*, **54**(4), 1508–22.

Stiglitz, J. (2000), "Two Principles for the Next Round or, How to Bring Developing Countries in from the Cold", *World Economy*, 437–54.
Telser, L. (1987), *A Theory of Efficient Cooperation and Competition*, Cambridge: Cambridge University Press.
Tichy, G. (2002), "What Do We Know About Success and Failure of Mergers?", *Journal of Industry, Competition and Trade*, **1**(4), 347–94.
Tybout, J. (2000), "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?", *Journal of Economic Literature*, **XXXVIII**, March, 11–44.
UNCTAD (1999), *Trade and Development Report*, Geneva: UNCTAD.
UNCTAD (2000), *World Investment Report*, Geneva: UNCTAD.
UNCTAD (2001), *World Investment Report*, Geneva: UNCTAD.
Vickers, J. (1994), *Concepts of Competition*, Inaugural Lecture as Drummond Professor of Political Economy before the University of Oxford, Oxford: Clarendon Press.
Wade, R. (1990), *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton, NJ: Princeton University Press.
Waring, G. (1996), "Industry Differences in the Persistence of Firm-Specific Returns", *American Economic Review*, **86**(5), 1253–65.
World Bank (1991), *World Development Report: The Challenges of Development*, New York: Oxford University Press.
World Bank (1993), *The East Asian Miracle: Economic Growth and Public Policy*, New York: Oxford University Press.
World Bank (2002), *World Development Report: Building Institutions for Markets*, New York: Oxford University Press.

13. Models of saving, income and the macroeconomics of developing countries in the Post-Keynesian tradition

Valpy FitzGerald

INTRODUCTION

The ex post resolution of ex ante imbalances between saving and investment intentions lies at the heart of the Keynesian approach to macroeconomics, and also – albeit in a longer-term context – to the 'classical tradition' in development economics associated with seminal authors such as Kalecki, Kaldor and Lewis. In this latter context, the focus is on capital accumulation and growth on the one hand, and the relationship between functional income distribution and the reinvestment of profits on the other – a focus which defines Thirlwall's own work in this field. His major contribution to modern development economics has been to preserve this classical tradition from the onslaught of neoclassical general equilibrium analysis and to show how Keynesian concepts still have empirical and policy relevance to industrialising countries. The key medium for communicating these ideas has probably been his textbook *Growth and Development*[1] which has run through seven editions since 1972 and remains today not only the leading introductory 'heterodox' development economics text but also the leading textbook for UK undergraduates in this subject.[2]

Thirlwall divides savings theory into three categories.[3] The first is the classical prior-savings model, under which he includes the current approaches of the World Bank and the International Monetary Fund, where saving is determined by income and financial incentives, and thus represents a constraint on investment (and thus growth) which can only be released by reduced fiscal deficits ('crowding out') or foreign capital inflows ('external saving'). The second is the Keynesian model, particularly the absolute income hypothesis: 'the Keynesian approach […] rejects the idea that savings determines investment and argues instead that the encouragement of investment will generate its own saving, either through increases in output if resources are unemployed, or through income

redistribution from groups with low propensity to save to groups with a high propensity to save as the result of inflation if resources are fully employed'. The third is what he calls 'quantity theory': a post-Keynesian model 'which emphasises the role of government monetary expansion in appropriating resources for development through forced saving or the inflation tax'.

In this chapter, I take up two of these models of saving behaviour in developing countries and try to show that they can be expressed in more formal terms following the familiar modern inter-temporal and open macro frameworks. In Part 2, I set out my understanding of his view on the relationship between income per capita and the aggregate savings rate; and show how this can be expressed formally as a wealth effect in the inter-temporal framework – which, in turn, allows some of his qualitative insights on financial liberalisation to be clarified. In Part 3 I look at his model of the link between wages, inflation and saving derived from Kaldor's distributional theory: this too can be placed within the modern standard macro-framework to some advantage in terms of clarifying the implications of the most important price in an open dependent economy – the exchange rate. Part 4 contains an extension of these ideas by including autonomous investment functions (which Thirlwall does not consider formally) to allow the analysis of the medium term effects of the real exchange rate on real wages and thus investment and employment – leading towards a post-Keynesian approach to the adjustment process. Part 5 concludes.

THE SAVINGS RATE AND PER CAPITA INCOME LEVELS

Inevitably, domestic saving must provide the greater part of the resources required for capital accumulation and economic growth. This includes savings by firms (that is, reinvested profits – including those of foreign firms) and governments (the fiscal balance *before* public investment) – as well as by households and individuals. It is clear from the evidence not only that the ratios of aggregate saving to national income in most developing countries are too low to sustain rates of investment and growth sufficient to rapidly reduce poverty, but also that this ratio rises with per capita income levels. Chapter XX of *Growth and Development*, entitled 'Financing Development from Domestic Resources', summarises Thirlwall's approach to this problem.

The Keynesian absolute income hypothesis is written (in Thirlwall's notation) with aggregate saving (S) as a function of aggregate income (Y):

$$S = a_0 + b_0 Y \tag{1}$$

which he divides through by population (N) to produce an expression for the level of savings per capita (S/N) of the form:[4]

$$S/N = a_1 + b_1(Y/N) \tag{2}$$

After dividing through by income per capita (Y/N) he obtains

$$S/N = b_1 - a_1(Y/N)^{-1} \tag{3a}$$

or in terms of income per capita (y) itself

$$S/N = b_1 - \frac{a_1}{y} \tag{3b}$$

and can thus conclude 'that the savings ratio (S/Y) is a hyperbolic function of the level of income per capita; that is, that the savings ratio will rise with the level of per capita income but at a decreasing rate' (*op. cit.* p. 491).

Thirlwall gives various possible explanations for this 'hyperbolic function' in developing countries: the increased monetisation of the economy in very underdeveloped countries logically raises the savings rate in the form of cash balances and bank deposits but this naturally reaches a saturation point when money demand is satisfied; the rising inequality of income in the early stages of industrialisation (following Kuznets) promotes saving due to differential savings rates between rich and poor, which then levels off as the industrial structure matures; a life-cycle effect exists when income per capita is rising and the demographic structure is changing; and the growing capitalist sector has an effect on the share of profits in national income leading to increased reinvestment by firms (as in Lewis). His own non-linear specification is also tested in Hussein and Thirlwall (1999), where it is found to be more significant than the standard linear specification in the literature. He also found that 'financial deepening matters', in the sense that there is a strong positive relation between the domestic savings ratio and that of the quasi-liquid liabilities of the banking system to GDP.[5] However, as Thirlwall himself admits,[6] there is a marked contrast between his results and those of larger studies such as Loayza et al (2000), which find that the high interest rates and increased private credit flows associated with financial liberalisation actually *reduce* the savings rate.

This approach may be contrasted with the standard textbook exposition savings determination.[7] The representative consumer (or household) maximises a time-separable utility function subject to an intertemporal budget constraint expressed in terms of wealth (w). Note that this constraint, assumed routinely in the literature, is equivalent to assuming perfect capital markets. In the two-period context this yields the familiar Euler equation in terms of marginal utilities – $u(c)$ – and the ratio of the pure time preference rate (l) and the world interest rate (r).

$$\max u(c_1, c_2) = u(c_1) + (1+\rho)^{-1} u(c_2)$$
$$\text{subject to } c_1 + (1+r)^{-1} c_2 = w = y_1 + (1+r)^{-1} y_2$$

which yields the result

$$\frac{u'(c_1)}{u'(c_2)} = \frac{1+r}{1+\rho} \tag{1.1}$$

which implies that a higher (exogenous) interest rate, r, will lower present relative to future income in the familiar way. This result can then extended to the standard infinite-horizon overlapping generations (OLG) framework without difficulty, and produces essentially the same result in a dynamic context.[8] Savings is then implicit in the optimisation process as simply the residual of consumption: once consumption choices are made, savings take place and are automatically translated into increased capital stock.

The 'law of motion' for this process at any one point of time is then

$$s_t = y_t + rw_t - c_t \tag{5}$$

$$w_t = w_{t-1} + s_t \tag{6}$$

The complications that arise in moving from the dynamic model to aggregate savings behaviour shed valuable light on the behaviour of this important variable in developing (and developed) economies. We can explore this with a simple isoelastic utility function

$$u'(c) = c^{-1/\sigma} \tag{7}$$

Substituting this into the two-period Euler equation and re-arranging gives us the 'consumption tilt' (θ) in a simple and tractable form:

$$\theta = \frac{c_1}{c_2} = \left[\frac{1+r}{1+\rho}\right]^{-\sigma} \tag{8}$$

which, as before, leads to an increased 'tilt' when the interest rate rises, as current consumption cuts are translated into improved future consumption.

However, once we derive the savings function explicitly we see that the savings function at any one point in time is rather complicated. In a two-period formulation there is only saving in Period 1. The consumption tilt (θ) comes from the Euler equation. Savings in that first period (S_1) are thus determined by not only the tilt itself but also current (y_1) *and future* (y_2) income:

$$c_1 = y_1 - s_1$$

$$c_2 = y_2 + s_1(1+r)$$

$$c_1/c_2 = \theta$$

$$s_1 = \frac{y_1 - \theta y_2}{1+\theta(1+r)} \tag{9}$$

From this we can derive the Keynesian function in Equation (3), but it is now more clearly related to future income in relation to the present, and thus to *expected income growth*.

$$S/Y = s_1/y_1 = \frac{1-\theta(y_1/y_2)^{-1}}{1+\theta(1+r)} \tag{10}$$

And we now have a more soundly founded interpretation of the coefficients in (3) in terms of the consumption tilt (θ) and the interest rate (r):

$$b_1 = \frac{1}{1+\theta(1+r)}$$

$$a_1 = \frac{\theta y_2}{1+\theta(1+r)} \tag{11}$$

From this it is clear that in addition to the relative price (tilt) effect:

> the net effect of r on s is ambiguous: an increase of r reduces θ and shifts the c_1/c_2 towards future consumption as we saw above; but the increase in r also leads directly to lower s because households need less wealth to secure future consumption ('wealth effect')
>
> the income effect of current income (y_1) is clearly positive, but that of future income (y_2) is negative – there is less reason to sacrifice current consumption. This can also reflect the permanent income hypothesis, with saving depending on the deviation of current income (y_1) from the long-term trend (y_2).

Note that the core model result does not necessarily mean that savings should be positive at all times; indeed an optimal solution is only possible if household borrowing is allowed at some point over the cycle. In terms of the two-period model above, obviously

$$\text{where } \theta y_2 > y_1 \quad \text{then } s_1 < 0 \tag{12}$$

Indeed, if households cannot access credit then they will 'over-save' to finance investment: in many developing countries consumer credit has been deliberately restricted by the monetary authorities in order to promote investment through bank lending. One of the undesirable consequences of financial liberalisation thus tends to be a sudden rise in consumer credit and thus a concomitant fall in *net* household savings – because even though household demand for money remains stable, the acquisition of new liabilities in order to move towards the desired intertemporal equilibrium (as derived above) reduces net household assets.

In addition, an exogenous change in the value of wealth (W) at any one point in time will also affect savings.[9] Developing countries suffer from these shocks, which tend to be proportionately larger and more frequent than is the case for developed countries. They include commodity price fluctuations, volatile global financial markets, sudden liberalisation processes, resource booms and armed conflicts. In consequence, savings rates not only vary widely between countries due to demographic, institutional and cultural factors (as Thirlwall points out) but also can change suddenly in any one country over time – a phenomenon that he does not address. The unexpected nature of these shocks means that uncertainty – in the Keynesian sense as distinct from risk – is central to an understanding of savings as well as investment behaviour in open dependent economies.

Uncertainty about future consumption reduces saving. We can see this formally by remembering that the utility function is concave, and so by Jensen's inequality, then

$$E_t\{u'(c_{t+1})\} < u'(E_t\{c_{t+1}\}) \tag{13}$$

In other words, expected utility from consumption is less than the utility of expected consumption. As the stochastic Euler function is

$$\frac{u'(c_t)}{E_t\{u'(c_{t+1})\}} = \frac{1+r}{1+\rho} \tag{14}$$

Then with uncertain future conditions $u'(c_t)$ will be lower at equilibrium, and thus c_t will be higher. In other words, saving will fall. However, uncertain future consumption is highly correlated with uncertain future income, leading to precautionary saving that depends on the variability of future output, and in particular on the third derivative of the utility function – $u'''(c_t)$ – which is of course positive for an isoelastic utility function such as that in (7) above.

Although this result may hold at the aggregate level, it should be remembered that the poorer households within developing countries undoubtedly regard

saving as 'buffer' against consumption falling below survival levels,[10] in which case we would expect asymmetric behaviour in response to income fluctuations (saving falling more on the downswing than it rises on the upswing). A similar phenomenon arises when the household can both save and invest in human capital through health or nutrition provision.[11]

Nonetheless, whatever the equity implications, income distribution in developing countries is such that the poor account for a very small proportion of aggregate saving. In consequence, the aggregate savings ratio will be correctly captured by the post-Keynesian model discussed above.

SAVING, INCOME DISTRIBUTION AND INFLATION

A Keynesian approach to the determination of the aggregate savings level in developing countries must start from the position that saving adjusts to investment, rather than the other way around. However, output in developing countries is not constrained by domestic demand, but rather by structural capacity constraints (such as limited food supply from peasant agriculture, as in Kalecki),[12] by declining terms of trade for primary commodities (as in Singer), barriers to world trade in their exports (as in Prebisch), or by limitations to import capacity (as in Lewis). In consequence, the ex-post adjustment of savings plans to investment levels cannot come about through changes in output but rather through inflation.

As Thirlwall points out, the basic Keynesian notion that investment determines saving forms the backbone of neo-Keynesian growth theory, as expounded by Robinson and Kaldor. He summarises his own view thus: 'The Keynesian approach to the financing of development by inflationary means stresses, first, that investment can generate its own saving by raising the level of income when the economy operates below capacity, and by redistributing income from wage earners with a low propensity to save to profit earners with a high propensity to save when the economy is working to full capacity. Second, inflation itself can encourage investment by raising the nominal rate of return on investment and reducing the real rate of interest. […] In Keynesian models, investment is not constrained by saving, but by the inflation rate willing to be tolerated by wage earners who have had their real wages cut'.[13]

He then goes on to provide his own version of this model: 'a higher level of investment can raise the rate of capital accumulation by raising the profit rate and share of saving in total income, subject, of course, to the inflation barrier. The mechanism that gives this result is rising prices relative to wages […] It is interesting to consider, using a model like Kaldor's, how much inflation is necessary to raise the savings ratio by a given amount'. He sets up the model as follows: 'Taking a savings function of the Kaldor type, $S = s_w W + s_r R$, let

$$S_t = s_w W_0 e^{wt} + s_r (Y_0 e^{pt} - W_0 e^{wt}) \tag{15}$$

where W_0 is the initial wage bill, Y_0 is the initial income level, w is the rate of growth of wages and p is the rate of growth of money income (= the rate of inflation). Dividing by $Y_0 e^{pt}$ to obtain an expression for the savings ratio at time t, and rearranging, gives

$$\left(\frac{S}{Y}\right)_t = \left(\frac{W_0}{Y_0}\right) e^{wt-pt}(s_w - s_r) + s_r \tag{16}$$

Now let $w = a + \alpha(p)$, where a is the rate of autonomous wage change and α is the wage-price coefficient. [The savings ratio equation] may then be written as

$$\left(\frac{S}{Y}\right)_t = \left(\frac{W_0}{Y_0}\right) e^{p(\alpha-1)t+at}(s_w - s_r) + s_r \tag{17}$$

[…] (the) savings ratio can then be seen to depend on four factors: the difference in the propensity to save out of wages and profits ($s_w - s_r$); the wage-price coefficient (α); labour's initial share of income (W_0/Y_0); and the rates of inflation and autonomous wage increases already prevailing'.[14]

He then observes that if there is full pass-through of prices to wages ($\alpha = 1$) and if the savings rates are equal ($s_w = s_r$) then inflation will not increase the savings rate; but that for reasonable parameter assumptions then quite low rates of inflation will increase the savings ratio substantially.[15] However, the dynamics of the wage-price reaction function are not explored either at the microeconomic level of the firm, nor in terms of macroeconomic disequilibrium – although he does suggest that inflation is a function of the gap between planned savings and planned investment.[16] One possible way of filling this gap in the post-Keynesian tradition, proposed by Kalecki,[17] would be to address the trade-off between wage bills and profits at the micro level, so that at the macro level the consequences will depend to a great extent upon the pricing behaviour of firms and trade union bargaining power.[18] With a degree of monopoly power sufficient to ensure a stable markup (π) on costs, then the price level (p) depends on the wage rate (w) and labour productivity (n) – the inverse of the labour input per unit of production. The real wage (ω) is thus given by these parameters too, independently of the rate of inflation.

$$p = (1+\pi)\frac{w}{n}$$

$$\omega = \frac{w}{p} = \frac{n}{1+\pi} \tag{18}$$

There is a large modern literature in the neo-Keynesian tradition on wages and inflation that is concerned with how wages adjust to prices. Generally these models take the form of forward or backward-looking wage contracts where output is a decreasing function of the real wage.[19] The backward-looking contract is often takes the form of 'wage indexing' in moderately inflationary conditions, and the forward looking one reflects 'wage claims' under high inflation.

To model these contracts in our case we should take into account all past (or future) prices, but for simplicity we use the derivative with respect to time, which yields the change in nominal wages as linear function of the wage-price gap and an adjustment parameter (β). For backward looking (*1*) and forward-looking (*2*) contracts we have

$$\dot{w}_1 = \beta_1 (w - p)$$

$$\dot{w}_2 = \beta_2 (p - w) \tag{19}$$

from which we can derive by some manipulation the result for the rate of change of real wages over time as a function of the real wage level itself (ω) and the rate of inflation (λ) under the two contracts.

$$\frac{\dot{\omega}}{\omega} \cong \frac{\dot{w}}{w} - \frac{\dot{p}}{p}$$

$$\frac{\dot{\omega}_1}{\omega_1} = -\beta\left(1 - \frac{1}{\omega_1}\right) - \lambda$$

$$\frac{\dot{\omega}_2}{\omega_2} = \beta_2\left(1 - \frac{1}{\omega_2}\right) - \lambda \tag{20}$$

The key result here is that both of these functions are 'explosive' and in fact the only steady state is where $\dot{\omega} = \dot{\lambda} = 0$ and, of course, goods market equilibrium. This reveals the weakness of the Thirlwall/Kaldor model of forced saving through the redistributive effects of inflation as a means of financing growth in any but the short run. It also helps explain the widespread recognition of the logic of stabilisation policies which require, in effect, a change in the nature of the wage contract (that is, reducing β sharply) or holding down prices artificially (e.g. by pegging the exchange rate or imposing temporary price controls).

Moreover, in the modern world of independent central banks and inflation targeting, rates of domestic inflation sufficient to force up the savings rate are not feasible, and in any case the uncertainty this causes seem to reduce savings

– or at best, force them abroad in the form of capital flight. [20] In contrast, the most important single price in the open dependent economy is of course the exchange rate. Inserting the real exchange rate into the post-Keynesian distribution model via real wage determination would allow an important analytical linkage to be established and help explain the political economy of adjustment.[21]

Consider an economy with two goods, traded and non-traded, and their respective prices, P^* in dollars and P in domestic prices. With E as nominal exchange rate then the real exchange rate (e) is defined as

$$e = \frac{EP^*}{P} \tag{21}$$

We define the real wage rate (ω) in terms of the nominal wage (w) and the consumer price level (p)

$$\omega = \frac{w}{p} \tag{22}$$

where the consumer price level (p) depends on the proportion (ϕ) of the domestic good (and thus the foreign good) in the consumption basket

$$p = \phi P + (1-\phi)EP^* \tag{23}$$

As Dornbusch (1986) does, we define the price of the non-traded domestic good (P) in terms of the a given ratio (γ) to the *nominal* wage rate (w)

$$P = \gamma w \tag{24}$$

We can now combine these, substituting (21) and (24) into (23) and then plugging the result into (22) so as to derive an expression that links the real exchange rate (e) and the real wage rate (ω)

$$\omega = [\gamma\{\phi + (1-\phi)e\}]^{-1} \tag{25a}$$

or

$$e = \frac{1}{1-\phi}\left\{\frac{1}{\gamma\omega} - \phi\right\} \tag{25b}$$

In other words, real currency depreciation (that is, e rising) *must* also mean a real wage reduction, and vice versa. No wonder that overvaluation (whether

from commodity price booms or pegged exchange rates) is politically popular, and under-valuation (especially large real devaluations associated with stabilisation policies) is unpopular! This provides a 'political economy' component to Thirlwall model of the balance of payments constraint on growth. However, more importantly in the context of this chapter, it is possible to rewrite the original Kaldorian savings function $S = s_w W + s_r R$ in terms of the real wage (ω), real output per worker (n) and the link with the real exchange rate (e) established in Equation 25:

$$S = s_w W + s_r R$$

$$\frac{S}{Y} = s_w \frac{W}{Y} + s_r \left(1 - \frac{W}{Y}\right) = s_r - (s_r - s_w)\frac{W}{Y}$$

$$\frac{W}{Y} = \frac{\omega}{n} = [n\gamma\{\phi + (1-\phi)e\}]^{-1}$$

$$\frac{S}{Y} = s_r - \frac{s_r - s_w}{n\gamma\{\phi + (1-\phi)e\}} \tag{26}$$

In other words, there does exist a clear 'post-Keynesian' link[22] between the real exchange rate and the savings rate, where a 'higher' value of e (i.e. depreciation) shifts income towards profits and thus raises the savings rate. This could form the basis for a fresh approach to adjustment policy, one step towards which I take in the next section of this chapter.

INVESTMENT, EMPLOYMENT AND ADJUSTMENT

Thirlwall reminds us of Keynes's fundamental proposition that savers and investors are separate

> With the existence of money and monetary assets, the act of saving becomes divorced from the act of investing. Those who want to do the investing may be different from those who want to do the saving, and the process of capital accumulation is likely to require financial and credit mechanisms to 'redistribute' resources from savers to investors. Indeed with a banking system with the power to create credit, investment can take place *without* prior saving through the process of borrowing. In other words, saving funds investment, but does not necessarily *finance* it. Investment generates its own saving through increases in output and profits. In fact, in the early stages of development, savings may not be the major barrier to capital formation but rather an unwillingness or inability to invest. (*Growth and Development*, p. 484, italics in original)

However, in the developing country context, the three 'representative agents' (households, firms and banks) cannot be simply be seen as savers, investors and (passive) intermediaries respectively (FitzGerald, 1995). First, households invest as well as save – in housing, farm improvement, children's education and so on; while firms save as well as invest – retained profits are their major source of finance; in addition firms and households hold financial assets and liabilities. Second, households and firms are highly heterogeneous: poor households have different savings behaviour from rich ones, due to lack of access to financial markets and the need to ensure essential consumption; while large firms operate under different objectives and constraints from small firms. Third, banks are not just passive 'intermediaries': their profit objectives lead to credit rationing of small firms and poor households; while their balance sheet positions have consequences for capital markets and exchange rates.

More importantly, it is probably essential to provide a more active role for investment than Thirlwall does. The neoclassical models assume full employment of labour and capital, the possibility of freely transferring *existing* capital and labour between sectors, and a production function where capital and labour can be smoothly substituted for one another *within* sectors. At most, 'sticky' labour markets are attributed to 'distortions' arising from labour legislation. Investment as such is not an issue, therefore. However, if output capacity cannot be changed without investment, then adjustment is not an instantaneous reallocation process at all, even though there is surplus labour available in developing countries. There is a close parallel here with trade theory, because the textbook framework[23] assumes that trade opening brings about relative price changes that then lead to similar inter-sectoral resource allocation shifts. This framework suggests that the factor rewards to the abundant factor of production will thereby rise – this often being taken to be unskilled labour, so income distribution improves. Further, it is assumed that the non-traded sector (particularly the protected import-substituting industry) is capital intensive and exports (particularly agricultural primary products) are labour intensive, so there is net employment creation from trade opening – and by extension macroeconomic adjustment.

However, modern trade theory – even in the neoclassical tradition[24] – indicates that a wide range of distributional results can occur. First, the abundant resource in many developing countries is natural resources (not only land but oil etc.) so that rents rise with trade liberalisation – or in our case, with real exchange rate depreciation and/or structural adjustment. The distributional result then depends on who owns the natural resources: peasants versus agribusiness; state oil companies versus multinational corporations etc. Second, the pattern of international demand (and by extension aggregate domestic consumption demand in our case) is for goods using skilled rather than unskilled labour, an effect reinforced by foreign investment. Thus skilled wages will rise more than

unskilled, and income dispersion increases. Third, the factor intensity (that is, technology) assumptions are not representative of many developing countries, where the non-traded sectors (construction, transport, commerce, cashcrops, government) are labour-intensive while exports (oil, mining, hotels, agribusiness etc) are capital-, skill- and technology-intensive. Thus employment will fall with structural adjustment unless investment rises sufficiently in the export sector.

We can explore the relationship between adjustment and investment in the open dependent economy[25] within the post-Keynesian tradition in the following way (FitzGerald and Perosino, 1999). Consider an economy with two sectors, exports (x) and home goods (h). Wage rates (w) are equal in the two sectors, each of which has given import coefficients (m) and unit labour requirements (l). Each sector produces to capacity, but there is surplus labour and total employment (L) is determined by output in each sector employment generated. Real profits (R) in the two sectors are (in terms of home goods prices) are then determined by the real wage (ω) and the real exchange rate (e):

$$R_x = (EP^*Q_x - wL_x - m_xQ_xEP^*)/P = Q_x[(1-m_x)e - \omega l_x]$$

$$R_h = (PQ_h - wL_h - mQ_hEP^*)/P = Q_h\{(1-m_he) - \omega l_h]$$

$$L = L_x + L_h = l_xQ_x + l_hQ_h \tag{27}$$

A change (we assume an increase here) in e raises real profits in the export sector and reduces them in the home sector. However, substitution is *not* smooth or even symmetric because both sectors are at full capacity. The speed of adjustment depends on the level of investment and the employment effect is ambiguous – it depends on the relative unit labour requirements (l) in each sector.[26] The two-period problem for firms in each sector is whether to invest or not. If they do not invest, then capacity (and thus production, Q) falls automatically by the depreciation rate (δ). The representative firm in the export sector must maximise the present value (V) of the firm[27]

$$V = R(Q_1) + \frac{R(Q_1)}{1+r} - Z(I_1)$$

$$I \geq 0 \tag{28}$$

where the investment cost (Z) is an increasing function of investment (here assumed to be of quadratic form) and includes the real exchange rate (e) as capital goods are imported

$$Z(I) = ezI^2 \tag{29}$$

Substituting (27) and (29) into (28) and differentiating V with respect to I gives the equilibrium condition where the value of the (export) firm is maximised

$$\frac{dV_x}{dI_x} = [e(1-m_x) - \omega l_x] - 2ezI_x(1+r) = 0$$

$$Q_2^x = Q_1^x(1-\delta) + I_x = Q_1^x(1-\delta) + \frac{e(1-m_x) - \omega l_x}{2ez(1+r)} \tag{30}$$

By construction, in our example of e rising (that is, depreciation) investment in the home goods sector does not take place as real profits have fallen, and as previous capital stock had been adjusted to equalise marginal returns to marginal costs, the extra increment of capital must be unprofitable.

$$Q_2^h = Q_1^h(1-\delta) \tag{31}$$

We can now derive the employment impact of the shift in the real exchange rate by substituting the employment functions in (27) into (30)

$$L_2^x - L_1^x = l_x(Q_2^x - Q_1^x) = l_x\left[-\delta Q_1^x + \frac{e(1-m_x) - \omega l_x}{2ez(1+r)}\right]$$

$$L_2^h - L_1^h = l_h(Q_2^h - Q_1^h) = l_h \delta Q_1^h$$

$$\Delta L = L_2 - L_1 = \frac{1}{2z(1+r)}\left[1 = m_x - \frac{\omega}{e} l_x\right] - \delta L_1 \tag{32}$$

This result has three interesting characteristics. First, there is no guarantee that employment will increase: this depends above all on the labour intensities (l) of each sector: clearly if the export sector is more labour intensive ($l_x > l_h$) this will happen, but if not, *the outcome is uncertain*. Second, the ratio of RWR to RER (ω/e) is clearly critical to the outcome – if real wages fall then employment will rise. However, this wage-employment 'tradeoff' is *not* the result of factor substitution, but rather of investment incentives. Third, an increase in the interest rate (r) – a common feature of IMF-inspired stabilisation programmes – will *reduce* the employment gain. This is the reverse of the factor substitution effect which textbook theory would predict.

CONCLUSION

It is unfortunate that Thirlwall's insights on aggregate savings behaviour in developing countries are not sufficiently well known by 'mainstream' development economists. This is partly a problem of 'modes of discourse' in the sense that although he presents his theoretical arguments with considerable analytical rigour and algebraic elegance, he does not place them within the conventionally accepted framework of modern inter-temporal macroeconomics. In this chapter I have attempted to rectify this neglect by exploring the link between post-Keynesian savings models and the modern 'open economy macro' framework, an exercise that can enrich both sides of the debate if not reconcile them.

This chapter has attempted to translate some of the central post-Keynesian propositions on savings behaviour in developing countries into the conventional macroeconomic framework of inter-temporal dynamics and open economy analytics. I have shown that this exercise permits a clearer view of the role of both wealth and income distribution in the savings and adjustment process. This has two advantages: on the one hand, to clarify the reasons for low and unstable savings rates; and on the other, to re-introduce political economy considerations to growth models.

I have also shown how the real exchange rate can be linked to the real wage rate, and thus to post-Keynesian concerns. This in turn allows the distributional implications of macroeconomic adjustment to be addressed. I have shown how employment, a key aspect of this topic, can be modelled and by doing so a more realistic alternative to the neo-classical notion of smooth and instantaneous 'resource reallocation' can be derived by looking at the sectoral investment decisions.

NOTES

1. All references in this chapter are to the Seventh Edition, referred to in the text as *G&D*. The preface confidently looks forward to an eighth edition in 2008.
2. And also, and possibly even more importantly in terms of international reach, for students on interdisciplinary 'development studies' courses.
3. *G&D* p. 490.
4. Quite why or how Thirlwall redefines the parameters from (1) to (2) is unclear: $b_1 = b_0$ seems otiose, while $a_1 = a_0/N$ seems inconsistent.
5. Although it is not clear which of the parameters in the theoretical model would be affected by this.
6. *G&D* p. 494.
7. See for instance Obstfeldt and Rogoff (1997), Chapter 1.
8. In developing countries we can take the real interest rate as exogenous: either because it is set by the government where the capital account is closed, or by international markets (due to the uncovered interest rate parity condition) where the capital account is open.
9. Empirical aspects are discussed in Agénor & Montiel (2000), Chapter 3.2.
10. Obsfeldt & Rogoff (1997, pp. 94–5).

11. Gersowitz (1988).
12. And thus developing economies are fully employed in the Keynesian sense of 'no spare capacity in the consumption goods industries' (*G&D* p. 490).
13. *G&D* pp. 525–6.
14. *G&D* pp. 529–31.
15. See also Thirlwall (1974).
16. Presumably this could also be expressed as the gap between planned output and planned expenditure.
17. See FitzGerald (1993) for a full exposition.
18. For a good discussion of 'trades union wage bargain' models, see Heijdra and van der Ploeg (2002), Chapter 8.
19. Agénor and Montiel (2000) Chapter 12.
20. See FitzGerald (forthcoming).
21. In Chapter 17 of *G&D* the real exchange rate is included in Thirlwall's well-known model of the external constraint on growth (pp. 687–90), although when discussing exchange rates and devaluation, he clearly believes that it is very difficult indeed to achieve adjustment through real depreciation and of no guaranteed positive effect on growth (pp. 693–5). However, he does not relate this insight back to income distribution or to savings.
22. Note that in this model the markup in the non-traded sector (γ) is fixed in Kaleckian fashion but the profit share (R/Y) is variable because the profit rate in the traded sector varies with the exchange rate. This point is taken up in the next section of the present chapter.
23. That is, the Heckscher-Ohlin-Samuelson (or 'HOS') model.
24. See, for instance, Wood (1994).
25. I am assuming here that there is a trade balance, so that external capital flows are not an issue. For an excellent inter-temporal macro framework for exploring the interaction of saving, investment and the current account, see Sen (1994).
26. Note that this model could easily be disaggregated to handle skilled/unskilled labour.
27. Heijdra and van der Ploeg (2002) Chapter 2.3.

REFERENCES

Agénor, P.R. and P.J. Montiel (2000), *Development Macroeconomics*, 2nd edn, Princeton, NJ: Princeton University Press.

Dornbusch, R. (1986), *Open Economy Macroeconomics*, New York: Basic Books.

Drezen, A. (2000), *Political Economy in Macroeconomics*, Princeton, NJ: Princeton University Press.

FitzGerald, V. (1993), *The Macroeconomics of Development Finance*, Basingstoke: Macmillan.

FitzGerald, V. (1995), 'Hamlet without the prince: structural adjustment, firm behaviour and private investment in semi-industrialized economies', in P. Arestis and V. Chick (eds), *Finance, Development and Structural Change*, Aldershot, UK and Brookfield, USA: Edward Elgar, pp. 27–450.

FitzGerald, V. (2003) *Global Markets and the Developing Economy*, Basingstoke: Palgrave.

FitzGerald, V. (2005) 'Monetary models and inflation targetting in emerging market economies' in P. Arestis, M. Baddeley and J. McCombie (eds), *The New Monetary Policy: Implications and Relevance*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

FitzGerald, V. and G. Perosino (1999), 'Trade liberalisation, employment and wages: a critical approach', in G. Barba-Navarretti, R. Faini and G. Zanalada (eds) *Labour Markets, Poverty and Development*, Oxford: Clarendon Press, pp. 34–62.

Gersowitz, M. (1988), 'Savings and development', in H. Chenery and T.N. Srinivasan (eds), *Handbook of Development Economics Vol I*, Amsterdam: North-Holland, pp. 381–424.

Heijdra, B.J. and F. van der Ploeg (2002), *The Foundations of Modern Macroeconomics*, Oxford: Oxford University Press.

Hussein, K. and A.P. Thirlwall (1999), 'Explaining differences in the domestic savings ratio across countries: a panel data study', *Journal of Development Studies*, **36**(1), October, 31–52.

Kalecki, M. (1976), *Essays on Developing Economies*, Brighton: Harvester.

Loayza, N., K. Schmidt-Hebbel and L. Serven (2000), 'What drives private saving across the world' *Review of Economics and Statistics*, May.

Obstfeldt, M. and K. Rogoff (1997), *Foundations of International Macroeconomics*, Cambridge, MA: MIT Press.

Sen, P. (1994), 'Savings, investment and the current account' in F. van der Ploeg (ed.), *Handbook of International Macroeconomics*, Oxford: Blackwell.

Thirlwall, A.P. (1974), *Inflation, Saving and Growth in Developing Countries*, London: Macmillan.

Thirlwall, A.P. (2003), *Growth and Development*, Seventh Edition, Basingstoke: Palgrave Macmillan.

Wood, A. (1994), *North–South trade, Employment and Inequality: Changing Fortunes in a Skill-Driven World*, Oxford: Clarendon Press.

14. The influence of Keynes on development economics

John Toye

INTRODUCTION

Development economics, in its modern form, was struggling to be born just as John Maynard Keynes, ill and exhausted by immense public responsibilities, approached his premature death in April 1946. This fact alone meant that Keynes himself exercised no direct influence on modern development economics. His significance for the new sub-discipline was nevertheless considerable, but it was wholly derived from indirect sources.

The first of these sources was the influence that he had exercised on the thinking of other economists who, after his death, did participate in the creation of development economics. Economists who fit this description included Joan Robinson and Austin Robinson: they were Keynes's disciples and they contributed, in their very contrasting styles, to the early construction of development economics (Harcourt, 1998: 367–77). James Meade also falls into this category. So does Hans Singer, who worked on employment issues before coming to development policy in the late 1940s (Singer, 1938 and 1939).

The second source of Keynes's influence was through his posthumous legend. The Keynes legend tended to affect especially younger economists who had not known Keynes personally, but nevertheless nurtured a view of him as a towering economist and a powerful role model of some sort – even when they rejected particular economic doctrines of his.

Ironically, what Keynes actually wrote about economic development failed to shape the sub-discipline at its birth, being either inaccessible or neglected. It has taken at least a generation since Keynes's death to construct the more detailed and nuanced view of Keynes that is now available. As part of the widespread scholarly effort to move beyond the Keynes legend, the eleven Keynes Seminars at the University of Kent played a notable role. Tony Thirlwall's contribution to the series as animator, organiser, participant, rapporteur and editor over twenty years was absolutely central to its success. It is an achievement for which he deserves the gratitude of Keynes scholars everywhere.

KEYNES'S INDIRECT INFLUENCE ON EARLY DEVELOPMENT ECONOMICS

Keynes's most notable economic doctrine was that the economy could be in equilibrium at less than full employment.[1] The implication of his new concept of an involuntary unemployment equilibrium was that resources were being wasted: what men and women did not produce today through enforced idleness would never be produced at all. To eliminate involuntary unemployment by government intervention was therefore to prevent the waste of resources and to allow their potential output to be made actual. In industrial societies, this waste was dramatically visible in the 1930s. It was there for all to see in the form of silent factories and mines, whose workers stood idly around outside the locked gates. The political dangers of this economic waste were also obvious, and not only to Keynes. For him, they were the motive to advocate policies to eliminate unemployment and to explain publicly how and why his policies would work.[2]

Joan Robinson, Keynes's junior collaborator in the making of *The General Theory of Employment* (1936), thought that the concept of full employment needed more careful definition, and she distinguished between open and disguised unemployment. Open unemployment affected wage workers, who were laid off by their employers when demand fell. She believed that another type of unemployment was often present among own-account workers, though it was disguised by the fact that they kept themselves occupied in various ways. Disguised unemployment was particularly prevalent in the agricultural sector when there was excess population on the land. Perhaps as a result of her stay in India in the years 1926–8, she argued that in the agricultural sector, and by extension in economies that are dominated by that sector, many people had occupations that contributed little to economic output, but this low productivity was not obvious because they appeared to be occupied.[3] Here, too, was a waste of resources, just the same as with open unemployment.

In the early 1940s, Britain's Royal Institute for International Affairs began studies to assist the post-war reconstruction of Europe. In the course of these, it was calculated that at least 20 to 25 per cent of the agricultural population of Eastern Europe was surplus, in the sense that they could leave the land without reducing agricultural output (Royal Institute of International Affairs, 1943). It was natural to believe that this surplus labour could be used for economic development if those in disguised unemployment could be absorbed into higher productivity occupations such as modern industries.

The Oxford Institute of Economics and Statistics undertook a similar project of reconstruction studies, using the talents of a number of scholars exiled from central Europe – including Michal Kalecki, E. F. Schumacher, Josef Steindl and Kurt Mandelbaum. As part of this project, Mandelbaum (1945: 2) produced a

pioneering projection model of an industrialisation process, a process that he justified in the following terms.

> An expanding population adds continually to the number of people who are forced to work on fragmented or overcrowded holdings or on inferior soil where their productivity is nil or almost nil (disguised unemployment). If these surplus workers were withdrawn from agriculture and absorbed into other occupations, farm output would not suffer, while the whole new output would be a net addition to the community's income. The economic case for the industrialisation of densely populated backward countries rests on this mass phenomenon of disguised rural unemployment.

It should be noted that Mandelbaum's intellectual formation had been at the Frankfurt Institute for Social Research, working on the agrarian question as formulated by Marx and Lenin. So it would be unwise to link the idea of disguised unemployment too exclusively with Keynesian macroeconomics (FitzGerald, 1990: 3–25). What we have is two different intellectual streams flowing into a single pool.

Nevertheless, when Paul Rosenstein-Rodan wrote in 1943 about the post-war development of South and South-eastern Europe, he did so in Keynesian language. He referred to these regions as 'international depressed areas', owing to the problem of agrarian excess population and its disguised unemployment. Like Keynes, he believed that the price mechanism could not be relied on to generate the required adjustments, and that, in such circumstances, governments could play a positive role in raising national welfare. Yet what Rosenstein-Rodan said that the government would have to do was much more than what Keynes had advocated in industrial countries. Factories would have to be built and financed; labour would have to be trained to work in them; markets would have to be found for the new output, without disrupting those of the foreigners who might be a source of external finance.

The chosen development strategy of Rosenstein-Rodan was 'the big push': substantial investment in a wide range of light industries that would create a demand for each other's output. This adaptation of Keynes's ideas indicated that, although demand was a problem in disturbed wartime conditions, big problems of economic development also existed on the supply side of the economy, the sound functioning of which Keynes felt able to take for granted. Thus right from the start, development economists had to consider how both demand and supply factors could be manipulated to realise the potential output of redundant agrarian labour.

Given the much more ambitious scope of the task, not to mention the disrupted wartime and post-war context, development economics at its birth was distinctly more *dirigiste* in spirit and practice than Keynesian macroeconomics ever was. While in principle private entrepreneurs could build and finance the missing factories, in practice the uncertainties, shocks and co-ordination prob-

lems that affected private investment decisions meant that government would have to lead, in the beginning at least, and the private sector would only be able to follow. Frank Burchardt, then Deputy Director of the Oxford Institute, expressed the contemporary view of the matter thus.

> For under-developed poor countries with surplus populations and a general shortage of social and private capital, different, and, in many respects, more formidable, problems arise. If a fairly rapid rate of industrialization is desired, a higher degree of direct controls and other methods, not unlike those employed in a war economy, will probably be needed (Burchardt, 1944: iv).

Yet, given the government's leadership role in relation to development investment, only part of which could be financed from abroad, there was another lesson that could be learned from Keynes. The policy lessons of his economics were not limited to devising policies for eliminating involuntary unemployment in a depression. The same underlying logic worked in reverse, bearing on the opposite problem of excess aggregate demand causing inflation in economies that were already operating at full capacity. His 1940 pamphlet *How to Pay for the War* had addressed this problem in the context of Britain. It envisaged an economy soon to reach full capacity as a result of growing expenditure on war, and it anticipated inflation unless methods of limiting consumption sufficiently could be found. As well as additional taxation, he recommended a form of forced saving called post-war credits, by which immediate compulsory reductions in income would be compensated by discretionary repayments later. If one thinks of development expenditure as similar to war expenditure – in that its benefits will accrue, if at all, in the future – Keynes's ideas on planned compulsory saving as an alternative to the inflation tax remain relevant.

In the event, however, Michal Kalecki was more influential than Keynes in shaping the theory of development finance. During the war, Kalecki had criticised Keynes's plans for war finance. He argued that compulsory saving could be offset by reductions in voluntary saving, and that those on low incomes would not be fully protected. Kalecki wanted a more extensive system of physical rationing of goods than the one that Keynes advocated (Moggridge, 1992: 630; Osiatynski, 1997: 3–37). Kalecki was concerned not merely to reduce the pressure of aggregate demand, but also to apply controls such as rationing and price fixing to maximise the release of raw materials and labour for the war effort. Other economists engaged in formulating methods of managing the British war economy tended to follow one or other philosophy – either that of Keynes or that of Kalecki. James Meade, who followed Keynes, was known as a 'thermostater' because he regarded the control of the pressure of aggregate demand as a sufficient instrument of economic policy in itself. Austin Robinson, Richard Kahn, Brian Reddaway and others at the Board of Trade followed Kalecki in calling for additional controls. Hence, they were known as 'Gosplanners', after

the Soviet planning body GOSPLAN (Toye, 2003: Chapter 8). The spirit of the Gosplanners, with their emphasis on directly attacking supply side problems, was the one that inspired development economics, and not the exclusive concern of some post-Keynesians with aggregate demand.

This was due to the work of Kalecki himself, who played a major role in transferring his approach to British war finance first of all into the economics of post-war reconstruction, and then into the analysis of the financing of development (Arndt, 1987: 124–6). He stressed that the control of inflation during the course of economic development was not a purely financial matter to be achieved by fiscal and financial devices. It required the resolution of disproportions in the real economy, and in particular overcoming the problem of an inelastic supply of food (Osiatynski, 1993: 23–44). This was inextricably bound up with the intractable socio-political problem of land reform. More generally, rigidities and bottlenecks on the supply side became recognised as a defining characteristic of a developing economy, and the major reason why orthodox policies of macroeconomic adjustment could not be expected to be effective. This 'structuralist' view dominated development economics up to and including the advent of structural adjustment policies in the early 1980s.

Although Kalecki was the dominant influence on the content of early development economics, Keynes was responsible for providing the statistical framework used by development planners. According to Donald Moggridge, 'finance ministers around the world still approach the problems of economic management in an analytical framework which, perhaps for want of a better word, commentators rightly call Keynesian' (Moggridge, 1993: vii). Hans Singer has said that 'the national income framework of analysis first developed by Colin Clark on the basis of the Keynesian concepts […] is surely the most striking preservation of the Keynesian heritage in current development planning' (Singer, 1987: 67–8). After Clark, James Meade and Richard Stone extended the estimation of national income into an interlocking set of national income, expenditure and production accounts. This tripartite system was then applied, on the initiative of Austin Robinson, to selected British colonies. Examples were the work of Phyllis Deane on national accounts for Northern Rhodesia (now Zambia) and that of Arthur Lewis for Jamaica (Thirlwall, 1987: 144–5). These exercises and others like them provided the statistical infrastructure within which post-independence development planning took place.

The development planning of the 1950s and 1960s, which was heavily macroeconomic in emphasis, thus combined a Kaleckian view of how the economy worked with a Keynesian statistical apparatus. This combination proved less incompatible than it might appear prima facie. Its objective was to identify the highest feasible rate of investment – assumed to generate additional income through a simple linear relation, the incremental capital-output ratio. The national accounts framework (plus an input–output table) ensured that for any

particular rate of income growth, all the components of income, expenditure and production could be projected consistently. The desired rate of growth then had to be selected. This choice depended on judgements of feasibility, and these judgements were informed by Kaleckian considerations. Thus practical development planners like Arthur Lewis emphasised the difficulty of rapidly expanding the supply of food, and warned of the inflationary consequences of attempting to expand investment faster than was consistent with the realistically likely growth rate of the economy's food supply (Lewis, 1966: 43–4, 53–4 and 154–5). Given a set of feasibility constraints, the maximum rate of growth can be calculated, and its components planned consistently. Thus the Keynesian accounting structure was flexible enough to be used for the macroeconomic planning of an economy with Kaleckian features.

Nevertheless, this style of development planning had plenty of critics among development economists. Gunnar Myrdal, for example, objected that it classified expenditure on education as consumption, rather than as an investment, which is how he believed that it should be treated in the context of development (Myrdal, 1968: 1916–7). Dudley Seers, an economic statistician who had worked for Kalecki at the UN, launched a more thorough attack, denying that achieving development was equivalent to maximising the rate of growth of the gross national product (Seers, 1969: 2–6 and 1976). He criticised the omission of economic activity in the informal sector, which he saw as a positive force for development, but he also objected to the neglect of income distribution, environmental costs and unpaid labour normally undertaken by women (Toye, 1989: 53–8). This critique gave considerable impetus to the search for alternative measures of development, such as UNDP's Human Development Index, which appeared in 1990. None of the alternatives has yet ousted the Keynesian framework from economic management in developing countries, which still survives in ghostly form in the World Bank's RMSM programming model.

From the mid-1970s, the Keynesian consensus in developed countries began to disintegrate. The experience of 'stagflation' – slower growth with accelerating inflation – could not be easily explained in terms of the Phillips Curve, according to which there was a growth-inflation trade-off. The division of economics into 'macro' and 'micro', for which Keynes was responsible, now began to appear problematic. Keynes had persisted with the microeconomic assumption of competitive markets, although it was not consistent with his theories of the behaviour of economic aggregates. The inconsistency raised the question of how to work out logically adequate micro-foundations for macroeconomics, leading to the sprouting of new less-than-Keynesian macro theories in the 1980s and 1990s. The search for the missing micro-foundations affected development economics adversely. Neither Kalecki nor Lewis had developed the economic theory of surplus labour and the agricultural marketed surplus sufficiently rigorously at the micro-level. Development economics was thus increasingly

criticised, as the star of Keynesian macroeconomics waned in developed countries.

THE LEGEND OF KEYNES

In addition to the indirect influence that Keynes exerted on specific elements of development economics through his disciples, he also affected development economists through his powerful posthumous legend. Indeed, he was already a legend in his own lifetime. It was a legend that he had cultivated, because he thought that it enhanced his power of persuasion in matters of public policy. After his death, Roy Harrod, who published the official biography of Keynes in 1951, further burnished this legend, again for the purpose of preserving the persuasive influence of Keynesian ideas (Harrod, 1951). Harrod's *Life* portrayed Keynes as 'a brilliant thinker, an entrancing personality and a great world benefactor'. Reviewers who did not know Keynes, or only did so in his final heroic years, found Harrod's portrait moving and convincing (Skidelsky, 2000: 494). Many younger development economists must also have been affected in this way. For them, Keynes was became icon and a role model.

What image of Keynes did the legend project? Keynes had shot to international fame at a stroke, when he published *The Economic Consequences of the Peace* (1919), his great polemic against the economic terms of the Treaty of Versailles. After that he was never long out of the public limelight as he published his policy analyses and advice in the press and corresponded with world leaders on economic issues. It was unprecedented for an economist to operate in this degree of publicity, even if policymakers frequently ignored his advice. During the inter-war years, he was as much a political figure in Britain as any of the elected politicians of the day. Originally regarded as rather unsound, he was later fully accepted by the political establishment, becoming the country's leading economic statesman. For a while, he seemed to hold Britain's post-war future in his hands. No economist had been in such a powerful position before.

The other aspect of the image was Keynes's intellectual stature, exhibited in his *General Theory of Employment, Interest and Money* (1978: VII). In this, he had provided a theoretical explanation for the persistence of mass unemployment and a theoretical justification for the policy of increased government expenditure in a depression – a policy that many other economists decried. He had shown the importance of analysing the aggregate demand and supply of the economy, the interdependence of monetary and real factors in inducing economic fluctuations and the inadequacy of looking only to price adjustment in the labour market to reduce unemployment. He created a new 'macroeconomics', very different from what had existed before – which might be described in

summary terms as Alfred Marshall's microeconomics plus the Cambridge version of the quantity theory of money. He saw himself, and other saw him, as creating a revolution in the discipline of economics.[4] The fact that the Keynesian revolution retained an unexplained separation of macro and microeconomics did not inhibit its initial success.

This extraordinary combination of political prominence and successful academic innovation won Keynes a unique place in British economics, as well as great respect among economists in the rest of the world. After his death, his legend became an ideal for economists to emulate, including many of those drawn into the professional world of development economics. To them, it seemed obvious to make a parallel between the scourge of mass unemployment in the 1930s and that of the persistence of global poverty (despite Western affluence) in the 1950s. The latter problem, like the former, seemed to require the economists who succeeded Keynes also to undertake a theoretical revolution, on the grounds that the economics of the day lacked the tools to provide an effective remedy.

The effect of the Keynes legend, then, was to encourage development economists to seek for theoretical revolutions. As the title of his masterpiece suggested, Keynes claimed to have created a more general theory of employment than the classical economists had. Neo-classical economists subsequently reversed this claim, however, arguing that Keynes had invented only a special case of neo-classical theory – the case where the wage level is fixed. Anti-Keynesian development economists on the political left, like Dudley Seers, inverted this argument as follows. All neo-classical economics, including the Keynesian variant, is a special case – the case of smoothly functioning markets in capitalist economies.

Thus, a new theory was sought in order to embrace the more general case, the case of the developing countries, whose economies are beset with bottlenecks and rigidities, and where well functioning markets are absent (Seers, 1963: 77–98). Albert Hirschman later changed the general/special distinction into a simple dualism. He argued that Keynes in his 1936 book had defeated 'monoeconomics'. His invention of another, radically different economics opened the way for the birth of development economics; 'the idea that there might be yet another economics had instant credibility'. In that sense, the step that Keynes had taken was 'crucial' for the new development economics (Hirschman, 1981: 6-9). One thing that unites Seers and Hirschman is that they argue – not without some paradox – from the irrelevance of Keynes's economics in developing countries to his significance for development economics. They also share a tendency to emphasise the absoluteness of the distinction between development economics and other branches of the discipline. In this they are following the Keynes legend that deliberately played up the fundamental nature of his quarrel with 'the classics'. However, just as Keynes exaggerated the depth

of his quarrel, so Seers and Hirschman are in danger of exaggerating the theoretical gulf between development economics and other sub-disciplines of economics.

The practical inheritance of the Keynes legend derived from the aftermath of his final efforts at designing world economic institutions that would provide for stability, growth and welfare. This was subsequently regarded in many quarters as unfinished business. Keynes's design for an international clearing union, and his disciple James Meade's design for an international commercial union, continued to command support long after they had lost out to US institutional blueprints for the new international institutions for finance and trade. Keynes's vision for the international economy was based on his long-standing opposition to deflation as a method of economic adjustment. He believed that disappointing the expectations of the rentier via inflation was preferable to disappointing the expectations of the worker via unemployment. He therefore wanted an IMF that was better resourced than the one that came about, and he wanted the burden of balance of payments adjustment to be shared by deficit and surplus countries, rather than being placed only on the former. The issue of whether the working of the International Monetary Fund could be adjusted in order to have a less deflationary impact on developing countries continued to be agitated through the 1950s and 1960s. So did the question of whether an International Trade Organization (ITO) (including its provisions for international commodity agreements) should supersede the purely interim GATT arrangements.

The claim – made by Gottfried Haberler and Raul Prebisch among many others – that the Bretton Woods system was incomplete, and moreover unhelpful to developing countries in relation to trade and finance, eventually prompted the birth of UNCTAD in 1964. Keynes had worked on a wartime proposal for a series of commodity buffer stocks (Keynes, 1980: 105–99). The policy of commodity buffer stocks remained popular with developing countries, and Gamani Corea made an integrated programme of buffer stock finance the central feature of the minimally successful UNCTAD campaign for a New International Economic Order from 1974 to 1980 (Toye and Toye, 2004: Chapter 10). The Report of the Independent Commission on International Development Issues (1980), chaired by Willi Brandt, was perhaps the last major incarnation of this particular Keynesian inheritance in the realm of international economic policy (Independent Commission on International Development Issues, 1980).

The attitudes created among development economists by the legend of Keynes are not to be taken wholly at face value. There was surely some element of disingenuousness in these claims, based on a Keynes-like appreciation of the value of hyperbole in getting oneself heard and understood. Even if Keynes did argue for a more generous design of the Bretton Woods institutions, for example, it would be an act of faith to believe that, had his view prevailed, his institutional

design would necessarily have lasted very long without subsequent adjustments (Williamson, 1987). We are here in the world of historical might-have-beens.

Certainly, Keynes's enthusiasm for commodity buffer stocks has been much exaggerated. His first proposal (1938) was made in anticipation of the outbreak of the Second World War, to avoid the difficulties of food imports experienced in the First (Salter, 1961: 255-9). His second proposal was intended to counter what he regarded as a much more damaging US plan for an international wheat cartel in 1942. When his own scheme for 'commod control' was rapidly and generally dismissed as impractical, because of the opportunities that it would have created for speculators, he did not return to the theme (Skidelsky, 2000: 234–6).

The legend of Keynes was significant for development economics because it constituted a standing temptation to its practitioners to indulge in the grandiose, both in the sphere of theory and of practice. The search for the next revolution in economic theory, or the next new paradigm, could easily degenerate into mere shifts of conceptual emphasis, which alienated more than they excited. In the end, the search for a new international economic order became an attempt to defy the laws of gravity in international political negotiation.

THE HISTORICAL KEYNES

The legend of Keynes thrived in the period when personal memories of him had faded or grown dim, and before a more authoritative historiography of his life and work produced its fruit. There were many elements that composed this constructive or re-constructive work. Archivists sorted and catalogued his papers. Scholars selected from them for publication of a substantial edition of his collected writings. Robert Skidelsky and Donald Moggridge, great authorities on Keynes, laboured to produce two excellent biographies.

During this time the Keynes Seminars in Canterbury made their own special contribution to the recovery process. The idea of bringing together members of the Keynes family, economists who had interacted with Keynes during his lifetime and contemporary analysts of the themes and policies that Keynes had made his own was an inspired one. Eleven published volumes, usually edited by Tony Thirlwall, survey almost every aspect of Keynes's career and intellectual interests. Less obviously, much sharing of information went on in the social and informal part of the seminar. Certainly, I can speak of the stimulus that I derived from my own participation to research Keynes's own forays into the area of economic development. These forays were directed towards two very important subjects, namely state planning and population. Yet Keynes's own views on these topics exerted little influence on development economists throughout the post-war period. In the early constructive phase of development

economics, the views of the historical Keynes on the economics of development remained neglected or unknown. The irony of it is that some of his views were produced again, as if new, by later commentators, including those that carried an anti-Keynesian banner (e.g., Johnson, 1978).

Keynes visited Soviet Russia three times – in 1925, 1928 and 1936 – and then wrote about what he saw. In his commentaries on the management of the Soviet economy, Keynes raised important issues that were to surface again in the neo-liberal critiques of development economics of the 1970s and 1980s.[5] His comments were not made initially from a position of outright opposition to the Bolshevik revolution: at first, he hoped that it might also embody 'some speck of the ideal'. He also advocated agricultural loans to the new regime, to prevent famine and promote the resumption of grain exports (Keynes, 1978: 394). It was only at the end of the 1920s that he abandoned these initial hopes.

Keynes analysed key features of the Soviet economy in his *A Short View of Russia* (1925). There, he described 'the official method of exploiting the peasants'. It was not done by taxation, but by the official manipulation of domestic prices that the foreign trade monopoly made possible (Keynes, 1978: 264). The peasants' wheat could be bought at a price in domestic currency below the world price at the official exchange rate, while textiles and manufactured goods would be priced in domestic currency above the world price at the official exchange rate. This price policy created the protection necessary for an internationally uncompetitive industrial sector. Keynes saw the ill consequences clearly enough: given the incentives he faced, the peasant would under-produce, and given the inflated industrial wage, there would be excess migration to the towns. Keynes summarised his analysis in the following terms.

> This state of affairs serves but to enforce a lesson of bourgeois economics as being equally applicable in a Communist state, namely that it impairs wealth to interfere with the normal levels of relative prices or with the normal levels of relative wages … (Keynes, 1978: 265).

The question must be asked whether Keynes's appeal to bourgeois economics was his last word on the matter. Surely, he later changed his mind and showed there was 'another economics', apart from the orthodoxy in which he had been schooled by Marshall? In the 1930s, did he not become an advocate of both protection and state planning? Indeed, he did, but it should be emphasised that these two shifts in his policy stance still left him at a considerable distance from the type of economic policies that he criticised in Russia in the 1920s. Even while making his cautious and limited case for protection in Britain, Keynes went out of his way to say: 'Russia exhibits the worst example which the world, perhaps, has ever seen of administrative incompetence and of the sacrifice of almost everything that makes life worth living to wooden heads'. He denied that by proposing a measure of protection for Britain, he was thereby endorsing 'all

those things which are being done in the political world today in the name of economic nationalism' (Keynes, 1982: 243–4). He contrasted the political context in which he was operating, one that allowed for experiment and free discussion of economic policy, with that of Stalin's Russia, where independent views were crushed.

Keynes also made his case for modest state planning by explicitly rejecting the relevance of what had been done by existing 'planned regimes to the south and the east' (Keynes, 1982: 86). These regimes, he believed, had bought their limited industrial successes at a huge cost in welfare. By contrast, his support for state planning was limited to activities 'which in the nature of the case lie outside the scope of the individual' (Keynes, 1982: 88). They were to be complementary to, and not a substitute for activities the individual could undertake. He was talking about measures of taxation, tariffs, foreign exchange control, the regulation of transport as well as town and country planning. He was not advocating a state-driven industrialisation drive. When he conceded that the ideas of the *General Theory* implied 'a large extension of the traditional functions of government', he also said that they gave no warrant for a 'system of State Socialism which would embrace most of the economic life of the community' (Keynes, 1978: 378).

The historical Keynes, while advocating greater control of investment by the liberal state, maintained, to the end of his life, the objections that he had voiced in the 1920s to the 'official method of exploiting the peasants'. If anything, his comments on the political economy of the Soviet Union became sharper and more hostile, as he became more aware of the enormity of the economic mistakes that were possible under authoritarian regimes.[6] As Joan Robinson said of him much later: 'Capitalism [was] in some ways repugnant to him but Stalinism was much worse' (Robinson, 1975).

The role of population in the process of economic development was also a matter of great interest to Keynes. From an analytical viewpoint, the most interesting aspect of his early work on population was a sketch of a model of the welfare implications of differential population growth in what would be known today as the North and the South. It is not fully worked out and abstracts from some crucial aspects of the situation that it addressed. What Keynes presented was essentially an example of an 'isolation paradox', a paradox of population growth that pre-figured his later famous paradox of thrift in the *General Theory.* As he put it, 'every patriot urges his country forward on a course of action [i.e. rapid reproduction] which is 'in the widest sense anti-social'. This pre-figures his later argument that the more virtuous people are in being thrifty, the further the national income will fall (Toye, 2000: 66). Such paradoxes, based on the insight that the structure of individual incentives is inconsistent with the achievement of the social good, have characterised political economy and economics from Mandeville to the modern environmentalist concern for the 'tragedy of the

commons'. They provide one link between the intellectual approaches of Keynes and modern development economists.

Surprisingly, Beveridge outwitted Keynes's in a debate on population issues in the 1920s, an event that partly explains his rather enigmatic essay of 1930, *Economic Possibilities for our Grandchildren*. In this, he deserts his characteristic preoccupation with the short run, and, for once, looks far into the future. He conjectures that within a hundred years, the economic problem will be solved, and the pressing problem will be the psychological one of how to adapt to the abundance of leisure. He celebrates the idea that by then, capitalism will have fulfilled its historic mission, liberating people to put aside the dubious arts of moneymaking and to attend to the more morally uplifting arts of life. Some have seen this as his acceptance of the stationary state of the classical political economists. A more likely explanation is that he was trying to induce some public optimism amid the gloom of the Great Crash, and that while doing so, he took the opportunity to make an unacknowledged recantation of his neo-Malthusian views.

In the 1930s, Keynes reacted to the population projections of Enid Charles and other British demographers. He understood them as a forecast that the British population was about to cease its long period of growth and to embark on a precipitous decline. His Galton Lecture of 1937 focused on the implications of population decline for effective demand, saving and the rate of interest. He concluded that excessive population decline would threaten to increase unemployment, just as excessive population growth would threaten to reduce average living standards. Keynes's final view of the need for a balanced relation between the growth of population and capital paved the way for the balanced growth theories of Roy Harrod and Evsey Domar in the 1940s.

The Harrod–Domar equation was useful in concentrating attention on the fundamentals of the development problem. It reinforced the perception that consistency was an essential requirement of the development planning process. The Harrod–Domar model has had a long life in the practice of development planning for a less creditable reason, however. Its doubtful assumptions of fixed coefficients between investment and growth, and foreign aid and investment and growth make it a simple, apparently objective yet misleading basis on which the international economic institutions can produce another set of investment and aid targets for their developing country clients (Easterly, 1999: 423–38).

Despite Keynes's recantation, the Malthusian threat still worried Keynes's disciples who wrote on the economic development of Asia. Austin Robinson reflected on the possibility of a Malthusian crisis in Bangladesh as late as 1974 (Robinson, 1974: 521–4). Joan Robinson courageously pointed out the dangers of population growth in the People's Republic of China, despite the Marxist doctrine that such statements were capitalist scare mongering (Harcourt, 1998: 374). 'Of all economic doctrines,' she asserted, 'the one most relevant to the

underdeveloped countries is that associated with Malthus' (Robinson, 1962: 107).

Were Keynes's views on population 'imperialist'? They did reflect one strand of 'imperial' discourse, one that was characteristic of certain late nineteenth century Indian civil servants. Some in the ICS favoured rural development projects, like canal building in Punjab, and the implementation of famine alleviation schemes. Others did not, believing that they were self-defeating in the face of steady population growth. The young Keynes belonged to the latter group (Toye, 2000: 104–5). Keynes also tended to think in race stereotypes, which combined biology and culture. It would be anachronistic to brand him as a racist, but he did hold caricature views of the national character of foreigners. These were widely shared in Bloomsbury and elsewhere, despite the fact that others at the time, like H.G. Wells, were perfectly well aware that they were false and unscientific.

CONCLUSION

The significance of Keynes for development economics is traditionally said to rest on two parts of his work. The first was his invention of the macroeconomics of employment, modulated by Joan Robinson into the concept of 'disguised unemployment'. The second was his contribution to the construction of the Bretton Woods international economic institutions, including his proposal for commodity price stabilisation. Both these components of his work were indeed influential in the early years of development economics, for two main reasons. Either they had influenced his disciples, who then wrote on economic development, or development economists who had not known him personally, perceived them as key elements in the legend of Keynes.

In contrast to these sources of influence, the beliefs of the historical Keynes about the subject of economic development had very little influence on its early formation. His neo-Malthusian beliefs led him (until 1930) to maintain that industry's terms of trade faced secular decline – the exact opposite of the Singer–Prebisch view that came to dominate development economics after 1950. In expositions of his neo-Malthusian approach, however, Keynes anticipated two modes of thinking that have subsequently become familiar in development economics, the isolation paradox and North–South modelling. His later concerns about the balance of investment and population change fed into the influential Harrod–Domar model.

The other significant theme to be found in the historical Keynes is his analysis of the use of economic controls in Soviet Russia in the 1920s and 1930s, when it was the contemporary equivalent of today's 'developing country'. In several ways, Keynes's perceptive critique of Soviet economic policy anticipated neo-

liberal criticisms of the industrial policies of many post-war developing countries. Ironically, some development economists of the 1970s, who proclaimed themselves anti-Keynesian because they were reacting against the legend of Keynes, were in fact elaborating on ideas that Keynes had sketched out fifty years previously, and of which they were blissfully ignorant. So easy it is, if the history of economics is disregarded or disparaged, to become 'the slave of some defunct economist'.

NOTES

1. By 'most notable', I mean the one that made the widest impact on the public consciousness, and not necessarily the one that the academics of the day found most original or controversial.
2. Keynes's political motivation is well captured in Harrod (1936).
3. The first reference to disguised unemployment dates from 1936, see Robinson 1936: 225–33, and recurs in her 'Planning Full Employment' (Robinson 1951: 84). See also her later 'Notes on Economic Development' (Robinson, 1960: 96–8).
4. 'I believe myself to be writing a book on economic theory, which will largely revolutionise – not, I suppose at once but in the course of the next ten years – the way the world thinks about economic problems', said Keynes in a letter to George Bernard Shaw of 1 January 1935. (Keynes 1982, Vol. XXVIII: 42).
5. I have treated this theme in greater depth in Toye (1993: 239–65).
6. Keynes's views thus contrasted with the distinctly more positive comments of Austin Robinson, one of the 'Gosplanners', made after visiting Soviet Russia in July 1945. I am grateful to Richard Toye for pointing this out to me. See Cairncross (1993:94) for more detail.

REFERENCES

Arndt, H.W. (1987), *Economic Development: the History of an Idea*, Chicago: University of Chicago Press.

Beveridge, William (1924), 'Mr Keynes' Evidence for Over-population', *Economica*, **4**: 1–20.

Burchardt, Fritz A. (1944), 'Foreword', in F.A. Burchardt, *et al.* (eds), *The Economics of Full Employment*, Oxford: Basil Blackwell.

Cairncross, Alec (1993), *Austin Robinson: the Life of an Economic Adviser*, London: St Martin's Press.

Easterly, William (1999), 'The Ghost of Financing Gap: Testing the Growth Model used in the International Financial Institutions', *Journal of Development Economics*, **60**: 423–38.

FitzGerald, E.V.K. (1990), 'Kurt Mandelbaum and the Classical Tradition in Development Theory', in Kurt Martin (ed.), *Strategies of Economic Development: Readings in the Political Economy of Industrialization*, Basingstoke: Macmillan.

Harcourt, G.C. (1998), 'Two Views on Development: Austin and Joan Robinson', *Cambridge Journal of Economics*, **22**: 367–77.

Harrod, Roy (1936), '[Review of] *The General Theory of Employment, Interest and Money*', *Political Quarterly*, **7**, April–June: 293–8.

Harrod, Roy (1951), *The Life of John Maynard Keynes*, London: Macmillan.

Hirschman, Albert (1981), *Essays in Trespassing: Economics to Politics and Beyond*, Cambridge: Cambridge University Press.
Independent Commission on International Development Issues (1980), *North–South: A Program for Survival*, Cambridge, MA: MIT Press.
Johnson, Harry G. (1978), 'Keynes and Development', in Elizabeth S. Johnson and Harry G. Johnson (eds), *The Shadow of Keynes: Understanding Keynes, Cambridge and Keynesian Economics*, Oxford: Basil Blackwell.
Keynes, J.M. (1978), *The Collected Writings of John Maynard Keynes*, Vol. VII, London: Macmillan.
Keynes, J.M. (1978), *The Collected Writings of John Maynard Keynes*, Vol. IX, London: Macmillan.
Keynes, J.M. (1978), *The Collected Writings of John Maynard Keynes*, Vol. XVIII, London: Macmillan.
Keynes, J.M. (1980), *The Collected Writings of John Maynard Keynes*, Vol. XXVII, London: Macmillan.
Keynes, J.M. (1982), *The Collected Writings of John Maynard Keynes*, Vol. XXI, London: Macmillan.
Keynes, J.M. (1982), *The Collected Writings of John Maynard Keynes*, Vol. XXVIII, London: Macmillan.
Keynes, J.M. (1983), *The Collected Writings of John Maynard Keynes*, Vol. XI, London: Macmillan.
Lewis, Arthur W. (1966), *Development Planning*, London: George Allen and Unwin.
Mandelbaum, Kurt (1945), *The Industrialisation of Backward Areas*, Oxford: Basil Blackwell.
Moggridge, Donald (1992), *Maynard Keynes: An Economist's Biography*, New York: Routledge.
Moggridge, Donald (1993), *Keynes*, Toronto: University of Toronto Press.
Myrdal, Gunnar (1968), *Asian Drama: An Enquiry into the Poverty of Nations*, London: Pelican.
Osiatynski, Jerzy (ed.) (1993), *Collected Works of Michal Kalecki*, Vol. V, Oxford: Clarendon Press.
Osiatynski, Jerzy (ed.) (1997), *Collected Works of Michal Kalecki*, Vol. VII, Oxford: Clarendon Press.
Robinson, Austin (1974), 'The Economic Development of Malthusia', *Modern Asian Studies*, **8**: 521–4.
Robinson, Joan (1936), 'Disguised Unemployment', *Economic Journal*, **XLVI** (182): 225–33.
Robinson, Joan (1951), 'Planning Full Employment', in Joan Robinson, *Collected Economic Papers*, Volume I, Oxford: Basil Blackwell.
Robinson, Joan (1960), 'Notes on Economic Development', in Joan Robinson, *Collected Economic Papers*, Volume II, Oxford: Basil Blackwell.
Robinson, Joan (1962), *Economic Philosophy*, Harmondsworth: Penguin.
Robinson, Joan (1975), 'What Has Become of the Keynesian Revolution?', in Milo Keynes (ed.), *Essays on John Maynard Keynes*, Cambridge: Cambridge University Press.
Royal Institute of International Affairs (1943), Committee on Reconstruction, 'Memorandum on Surplus Agricultural Population', London.
Salter, A. (1961), *Memoirs of a Public Servant*, London: Faber.
Seers, Dudley (1963), 'The Limitations of the Special Case', *Bulletin of the Oxford Institute of Economics and Statistics*, **25**(2): 77–98.

Seers, Dudley (1969), 'The Meaning of Development', *International Development Review*, **11**(4): 2–6.
Seers, Dudley (1976), 'The Political Economy of National Accounting', in Alec Cairncross and Mohinder Puri (eds), *Employment, Income Distribution and Development Strategy: Essays in Honour of Hans W. Singer*, London: Macmillan.
Singer, H.W. (1938), 'The Process of Unemployment in the Depressed Areas, 1935–38', *Review of Economic Studies*, **6**: 177–88.
Singer, H.W. (1939), 'The Process of Unemployment and Regional Labour Markets', *Review of Economic Studies*, **7**: 42–58.
Singer, H.W (1987), 'What Keynes and Keynesianism Can Teach Us about Less Developed Countries', in A.P. Thirlwall (ed.), *Keynes and Economic Development*, London: Macmillan.
Skidelsky, Robert (2000), *John Maynard Keynes, Volume 3: Fighting for Britain 1937–46*, London: Macmillan.
Thirlwall, A.P. (ed.) (1987), *Keynes and Economic Development*, London: Macmillan.
Toye, John (1989), 'Nationalism and Structuralism: Two Themes in the Work of Dudley Seers', *IDS Bulletin*, **20**(3): 53–8.
Toye, John. (1993), 'Keynes, Russia and the State in Developing Countries', in Derek Crabtree and A.P. Thirlwall (eds), *Keynes and the Role of the State*, Basingstoke: Macmillan, pp. 239–65.
Toye, John (2000), *Keynes on Population*, Oxford: Oxford University Press.
Toye, John and Richard Toye (2004), *The UN and Global Political Economy: Trade, Finance and Development*, Bloomington: Indiana University Press.
Toye, Richard (2003), *The Labour Party and the Planned Economy, 1931–1951*, London: Royal Historical Society.
Williamson, John (1987), 'Bancor and the Developing Countries: How Much Difference Would It Have Made?', in A.P. Thirlwall (ed.), *Keynes and Economic Development*, London: Macmillan.

Index