INTRODUCTION TO THE SERIES

The aim of the *Handbooks in Economics* series is to produce Handbooks for various branches of economics, each of which is a definitive source, reference, and teaching supplement for use by professional researchers and advanced graduate students. Each Handbook provides self-contained surveys of the current state of a branch of economics in the form of chapters prepared by leading specialists on various aspects of this branch of economics. These surveys summarize not only received results but also newer developments, from recent journal articles and discussion papers. Some original material is also included, but the main goal is to provide comprehensive and accessible surveys. The Handbooks are intended to provide not only useful reference volumes for professional collections but also possible supplementary readings for advanced courses for graduate students in economics.

CONTENTS OF THE HANDBOOK

VOLUME I

PART 3 – THE ECONOMICS OF ENVIRONMENTAL POLICY

VOLUME II

PART 4 – THE ECONOMICS OF RENEWABLE RESOURCE USE

PART 5 – THE ECONOMICS OF PROVIDING RENEWABLE RESOURCE GOODS AND SERVICES

PART 6 – ENVIRONMENT AND RENEWABLE RESOURCES IN SOCIALIST SYSTEMS

VOLUME III

PART 1 – SOME BASIC CONCEPTS

PART 2 – ANALYTICAL TOOLS

PART 3 – APPLICATIONS TO POLICY AND FORECASTING ISSUES

PREFACE TO THE HANDBOOK

Natural resources have been studied by economists from the earliest days of the profession. They have been seen as providing a basis for national prosperity, power, and wealth. The ability to harness energy in new ways has been recognized as a major, if not the major, factor underlying the industrial revolution. Because forests, fisheries, and agricultural land are fundamental to food supplies, these resources have been long studied.

Yet only relatively recently have there been developed broad theories specific to the fields of natural resources and energy economics. Previously, examination of these fields relied upon the general economic theories being utilized for analysis of other commodities. More recently, however, it has been recognized by economists that certain special characteristics of natural resources have required theories which explicitly accounted for these characteristics.

Agricultural land, forest, and fisheries have been seen only in the last generation to be usefully described as renewable resources. Such resources are self-renewing at a limited rate which may itself depend upon the size of the stock in existence at any given time and upon the extent and nature of human intervention into the stock dynamics.

Minerals and many energy commodities are now seen as depletable or nonrenewable resources. These are resources for which only a limited concentrated stock exists for allocation over all time. For these resources, a central issue involves *when* they should be extracted, since a decision to utilize a given portion of the stock at one moment of time precludes the opportunity of using that portion at another time.

Even more recently have the environmental resources – air, water, open space – been also seen as renewable or even in some cases depletable resources. The image of environmental resources, fisheries, and wild animal stocks as common property resources owned by everyone and hence by no one is also of relatively recent development. And even more recently, economists have systematically incorporated concepts of materials balance into theories of the flow of physical materials from the natural environment, through the economy, and back into the natural environment.

And it has been only since the early 1970s that energy resources have been given particular attention as a matter for theorizing, empirical testing, and policy-making.

Thus, there now exists a set of concepts which unite the field of natural resource economics. While these concepts are also finding application in other branches of economics, their formalization has been motivated by the need to better understand natural resource issues.

Also uniting the study of natural resource issues is the growing realization that most important energy and natural resource issue are inherently interdisciplinary. The interdisciplinary nature requires applied work to integrate information from some combination of physics, engineering, chemistry, biology, ecology, political science, and law.

To a lesser extent the current theories also reflect this interdisciplinary reality. Materials balance concepts from physics are now fundamental to economic theories of the environment. Population dynamics concepts from biology and ecology are intertwined with economic concepts in renewable resources theories. Thermodynamic concepts and concepts of energy conservation are fundamental to theoretical work on energy economics. Legal concepts of property rights and ownership greatly influence analysis of environmental economics.

The study of resource economics has thus required and motivated researchers to reach out beyond their own disciplines and to integrate ideas from other fields into their own disciplines. Presumably this integration will influence not only resource economics but also other areas within economics.

The three volume comprising the *Handbook of Natural Resource and Energy Economics* examine the current theory and sample current application methods for natural resource and energy economics. Volumes I and II deal with the economics of environmental and renewable resources. Volume III, which is still in preparation and whose outline is included in this volume, will deal with the economics of energy and minerals.

Volumes I and II are divided into six parts. Part 1, which deals with basic concepts, consists of five chapters. The first chapter discusses environmental issues and welfare economics. Among the more penetrating developments in the short history of environmental economics is a wedding of the concepts of economic general equilibrium, materials balance, and common property resources into a single unified theory. This model offers a systematic explanation of the occurrence of pollution-type environmental problems and an opportunity to explore the welfare economics of suggested remedies. In Chapter 1, Karl-Göran Mäler uses a version of this model to provide a general theoretical framework for the field of environmental economics.

Chapter 2 attests to the interdisciplinary character of both environmental and renewable resource economics. In it James Wilen explains the bioeconomic models pertinent to these fields. The response of biological systems both to insults and to management actions is a central concern in many natural resource problems. Often, models simulating these responses are an integral part of the economic analysis of such problems.

In much of economics the spatial relationships among economic activities can be safely ignored. In environmental economics these relationships can rarely be ignored. Environmental effects of human action occur in and through space; neglect of this fact can lead to serious error. Space is involved in such matters as the degradation of residuals in the environment, the effects of airborne residuals on visibility, and the efficiency of alternative environmental policies. Moreover, environmental economics must address problems of interregional and international trade. In Chapter 3, Horst Siebert explores the spatial aspects of environmental economics.

Conservation of natural resources is a long-standing human concern. But in the last two decades there has been active economics research addressing the problems related not to scarcity of resource commodities, but rather to the protection of natural areas. This research has concerned itself with such issues as irreversibility, option values, and asymmetric technological change. In Chapter 4, Anthony Fisher and John Krutilla address these new conservation issues.

The final chapter in Part 1 deals with ethics and environmental economics. The theoretical underpinning of benefit–cost analysis, one of the basic tools of natural resource economics, is welfare economics. Welfare economics, in turn, can be viewed as an enormous elaboration and adaptation of an ethical theory: classical utilitarianism. But there are other valid ethical systems. And these other systems might imply quite different outcomes if applied to natural resources problems. For example, issues such as the long-term storage of nuclear waste and changes in climate resulting from resource use raise ethical issues perhaps more strongly than is usual in economics. These concerns are addressed in Chapter 5 by William Schulze and Allen Kneese.

Part 2 deals with methods and applications of economics to environmental problems. In Chapter 6, A. Myrick Freeman reviews methods for assessing the benefits of environmental programs. One of the most challenging areas of environmental economics, development of methods for estimating benefits of environmental improvements, has also been one of the most active areas of research in recent years. The interest results, in part at least, from increased pressure to demonstrate benefits from the costly environmental improvement and protection programs put into place by governments of industrialized countries in recent years.

Another major area of environmental economics, pursued especially actively in the 1970s, is the application of quantitative (usually linear) economic models to environmental questions. Such models have been applied to analyze effects of alternative policies on residuals generation and on control cost at both the industrial and regional level of detail. For regional analysis transfer functions which translate emissions at various points into ambient concentration at other receptor points – are often embedded directly into economic models. David James reviews both industrial and regional models and their applications in Chapter 7.

An important class of linear models applied to environmental problems is that of national input–output models. When outfitted with residuals generation coefficients and residuals control options such models can be utilized to analyze indirect, as well as direct, effects on the environment of economic growth, changes in product mix, and alteration of other variables of interest. In Chapter 8, Finn Førsund describes the use of national input–output models, with special application to the economy of Norway.

Part 3 of the Handbook includes two chapters on the economics of environmental policy. Chapter 9, by Gregory Christainsen and Tom Tietenberg, reviews what is known about the distributional and macroeconomic consequences of environmental policy. How, if at all, does environmental policy contribute to inflation or to unemployment? How are the costs and benefits of environmental policy distributed among income groups? This chapter describes methods of addressing such questions and offers a set of conclusions.

Chapter 10, by Peter Bohm and Clifford Russell, provides a comparative analysis of environmental policy instruments. While the idea of effluent fees as a policy instrument flows naturally from abstract economic reasoning, most governments have chosen not to follow economists' advice and have resorted to command and control strategies. Also advocated by some economists, and partially implemented, are tradeable permits to emit residuals. Deposit-and-return systems are also applied to some environmental problems and may have potential for dealing with others. This chapter reviews what the last twenty years of economic research have shown about the strength and weaknesses of these various approaches.

Part 4 deals with uses of renewable resources other than simply as recipients of residuals. Water resource development and use has probably received more attention from economists than any other natural resources subject except agriculture. There are at least three reasons for this attention. Because federal water resources agencies have long practiced benefit–cost analysis in the evaluation of water resources, there has been much opportunity for economists to develop and use theoretical concepts, methods, and data for such evaluations. Second, the development of river systems for multiple purposes has provided interesting opportunities for the application of systems analysis, that close relative of microeconomics. Third, market processes have played some role in the allocation of scarce western water. Chapter 11, by Robert Young and Robert Haveman, reviews economic and institutional aspects of water development.

The remaining two chapters in this part, Chapter 12 by Michael Bowes and John Krutilla, and Chapter 13 by Alan Randall and Emery Castle, deal with land use, although not in the traditional manner as a factor of production in agriculture or yielder of a single product, wood, in forestry.

Chapter 12 deals with the management of wildlands. Recognizing that wildlands yield not only timber but also recreational and aesthetic values, this chapter integrates theory derived from the forestry literature with that from the multipurpose firm literature. Chapter 13 also departs from the conventional view of land, using an asset pricing model to analyze land markets. The chapter includes an in-depth study of rent determination, examining influences of macroeconomic changes and of growing alternative demand for land on land prices, and in turn examines the reaction of land prices to increasing rents. The chapter also explores implications for land use planning and regulation and examines the role of land in the evolution of economic thinking.

Part 5 deals with the economics of renewable resource goods or services provision. Chapter 14, by Anthony Scott and Gordon Munro, treats commercial fishery economics. Commercial fishing has fascinated natural resources economists because this activity uses a common property resource as an essential input. The common property nature of the resource in a free market leads to decisions which produce economic inefficiency. Free access can lead to excessive depletion of the resource and to excess investment, both phenomena eliminating any net economic returns that would, under optimal management, be available from this resource. The chapter reviews these issues and spells out implications for public policy and international cooperation.

Chapter 15, the final one in this part, by Kenneth McConnell, treats the economics of outdoor recreation. It surveys conceptual and empirical approaches, problems, and solutions encountered in applying economics to the provision of natural resources for recreational purposes. It also shows how the evolution of the economics of outdoor recreation was influenced by the distinctive nature of markets for outdoor recreation.

Part 6 concludes Volumes I and II with two case studies dealing with environment and renewable resources in socialist systems. The first, by Marshall Goldman, focuses upon the Soviet Union, and the second, dealing with China, is by Shigeto Tsuru.

Since in socialist states all means of production are owned by the state, a superficial view might suggest that all externalities would be internalized and that, therefore, there would be no incentive to generate excessive residuals or overuse renewable resources. Goldman, in his study, shows that for the Soviet Union this impression is very far from the truth. He argues that the incentives for abusing resources are at least as large as in market economies and, possibly, much larger. Tsuru's study of China suggests that the situation may be somewhat different there. China is a developing economy and resources for environmental protection are accordingly limited. There is, however, explicit recognition of the environmental problem, and there is a public policy aimed at the comprehensive recycling of wastes. Presumably, this recycling is motivated by the scarcity of resource inputs as well as by a desire for control of residuals.

ALLEN V. KNEESE
Resources for the Future, Inc.

JAMES L. SWEENEY
Stanford University

Chapter 11

ECONOMICS OF WATER RESOURCES: A SURVEY

ROBERT A. YOUNG

Colorado State University

and

ROBERT H. HAVEMAN*

University of Wisconsin

> Water is far from a simple commodity,
> Water's a sociological oddity,
> Water's a pasture for science to forage in.
> Water's a mark of our dubious origin.
> Water's a link with a distant futurity,
> Water's a symbol of ritual purity,
> Water is politics, water's religion,
> Water is just about anyone's pigeon.
> Water is frightening, water's endearing,
> Water's a lot more than mere engineering.
> Water is tragical, water is comical,
> Water is far from the Pure Economical.
> So studies of water, though free from aridity,
> Are apt to produce a good deal of turbidity.
>
> Kenneth Boulding (1964)

1. Introduction and overview

This chapter reviews the application of economic concepts to the study of the consumption, supply, and allocation of water resources. Water management poses a wide array of issues for the economist, since few commodities are so pervasively involved in human economic activities. To an important degree, the location and intensity of economic activities depends on the availability of water for drinking, for agricultural and industrial production, for sanitation and waste assimilation, for transportation and for aesthetic and recreational benefits.

Water is said to be the only substance which exists in all three physical states – solid, liquid, gas – within the normal temperature range found on the earth's surface. Via the process known as the hydrologic cycle, the earth's water

---

* We wish to thank H.C. Cochrane, S.L. Gray, M.L. Livingston, J. McColl, R.G. Walsh, and of course, the editor, for helpful comments on earlier drafts.

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

inventory is continually being transformed among the three states. No form of life on earth can exist without water. Water is a nearly universal solvent.

Enormous quantities of water are available; the earth's estimated stock exceeds tens of trillions of gallons per capita. Although only a tiny fraction of this amount is readily usable by humans – because most is too salty, in frozen or vaporous form, or simply in the wrong place – the world's fresh water supply is plentiful relative to present consumption patterns [Baumgartner and Reichel (1975)]. A "water problem" exists when water is not found in the proper quantity and quality at the appropriate place and time.

Objective and scope

Our aim is to direct attention to the more significant of the economic aspects of water resource management. Due to space limitations, we will concentrate on the approaches to policy evaluation, including both project appraisal and the assessment of incentive structures for water users. The emphasis is on the U.S. experience. Matters dealing with water quality and recreation are treated elsewhere in this volume, and are largely ignored here.

Section 2 reviews those characteristics of water resource systems that serve to set them apart from other resources, with particular reference to the attributes which serve as the basis for public intervention. It also describes the nature of the interventions which have been made, and emphasizes the need for evaluating them in terms of their objectives. This is followed, in Section 3, by a survey of cost–benefit concepts and procedures as they have been applied in the water resource planning area. Special attention is given to the measurement of economic benefits. The remaining sections survey several important policy issues relating to water allocation and development, including irrigation planning, floodplain management, interbasin transfers, pricing and allocative institutions.

2. Characteristics of water resource systems and patterns of supply and use

This section treats a number of specific characteristics of water and its use which are relevant to the economics of water and public intervention into water allocation. It also surveys water supply and water use patterns. [See also, U.S. Water Resources Council (1978).]

2.1. Water supply and use

Fresh water for human use may be found in *surface* water (open bodies of water such as streams or lakes) or *groundwater* (from subsurface zones in which water is found in voids in sands, gravels, etc.). Water generally is categorized among the renewable (flow) resources, although certain groundwater deposits are more usefully analyzed with concepts applicable to the non-renewable (stock) resource case.

The unique characteristics of water consumption mentioned above necessitates particular care in understanding what precisely is meant by water "use". Conventional terminology distinguishes between *offstream* and *instream* uses [Solley, Chase, and Mann (1983)]. Offstream uses are those requiring withdrawal or diversion from a ground or surface water source. Examples include crop irrigation, industrial water use for cooling or cleaning, and municipal water supply for consumption, cleaning and waste removal. Several factors are involved in measuring the amount of water "used" in an off-stream activity. *Withdrawal* refers to the amount of water diverted or pumped from the source of supply. *Delivery* means the amount of water received at the point of use, while *release* is the amount returned to the hydrologic system from the point of use. With *consumptive use*, water is no longer available because it has been evaporated, transpired, incorporated into products, or otherwise removed from the water environment. *Return flow* is that amount that reaches a ground or surface water source after release and thus becomes available for further use. *Conveyance losses* are waters lost in transit from pipe, canal, or other conduit by leakage, seepage, or evaporation. In certain cases, losses may be available for reuse, in which case they may be included as return flows.

Generally speaking, consumptive use plus conveyance losses plus return flows sums to withdrawal. Withdrawal and consumption are the two principal concepts by which water "use" is measured. However, use categories differ greatly in the quantity and quality of their return flows, and hence on the further usability of the non-consumed portion. A full evaluation of water use, therefore, must consider both quantity and quality dimensions.

Non-withdrawal (*instream*) *uses* are those uses requiring no diversions from ground or surface water sources. Examples include hydroelectric power generation, maintenance of streamflow or water supplies to support fish and wildlife habitat or aesthetic values, dilution of wastewaters, freshwater dilution of saline water bodies, and right-of-way provision for inland waterways navigation. A number of unresolved conceptual difficulties remain in quantitatively measuring non-withdrawal uses since the waters are neither withdrawn nor consumed. Those issues arise mainly in cases where the tradeoffs between instream and offstream uses are being assessed.

Table 11.1 summarizes estimates of water withdrawals and consumption for the United States in 1980. The major withdrawals of water are for industrial and irrigation uses, accounting for 51 percent and 40 percent, respectively. Since most industrial use is for thermoelectric power plant cooling, which is relatively non-consumptive, this category accounts for only 8 percent of national consumptive use. Irrigation water, which is about 55 percent consumed, accounts for a dominant 82 percent of total water consumption.

Table 11.1
Withdrawal and consumption of fresh water in the United States, 1980
(by source and category of use).

	Withdrawals (millions of gallons per day)			
	Groundwater	Surface water	Total	Consumptive use
Irrigation	60 000	90 000	150 000	83 000
Self-supplied industrial	11 600	179 000	191 000	8 200
Rural use[a]	4 400	1 200	5 600	3 900
Public supplied[b]	12 000	22 000	34 000	7 100
Total	88 000	290 000	378 000[c]	102 000

[a] Rural use includes domestic and livestock uses.

[b] Public supply is water withdrawn for all other uses by public and private water suppliers.

[c] Excludes 171 000 mgd of saline water withdrawn primarily for thermoelectric power plant cooling.

Source: Solley, W.B., E.B. Chase, and W.B. Mann (1983) *Estimated Use of Water in the United States*, U.S. Geological Survey Circular 1001.

Consumption patterns in other countries will, of course, vary by climate and degree of development. Irrigation represents the major consumptive use of water in the world, as in the United States.

2.2. Characteristics of water resources: The rationale for intervention

The logic of economics emphasizes private resource allocation decisions if the conditions required for a smoothly functioning market system exist. These conditions involve both the nature of goods being traded and the characteristics of the markets within which the trades occur. In brief, these conditions are that there must be perfect competition in the private factor and product markets. Competition, in turn, requires that: (1) Each industry in the economy exhibits increasing costs; (2) all goods and services produced and traded must be exclusive; (3) goods which exhibit jointness in supply, such that one individual's consumption does not diminish any other individual's use of the good (public goods) are absent; (4) all buyers and sellers must have full knowledge of all the alternatives available to them and the characteristics of these alternatives; (5) all resources must be completely mobile; and (6) ownership rights are clearly attached to all goods and services to be traded in the economy.

Physical and economic attributes of the water resource

On several scores, either water as a commodity or the markets in which water is actually bought and sold fail to meet the requirements listed above. In fact, markets in water are "rudimentary" and unorganized in that there is no regularity of procedure, intermediaries, or location [Brown et al. (1982)]. Several factors account for this situation. Some of these are related to the difficulties in defining "water use", as discussed above. Extending and modifying Bower's (1963) approach, some of the primary characteristics of water which account for the inadequacy of water markets can be listed as follows.

Mobility – Water tends to flow, evaporate, seep, and transpire. These attributes present problems in identifying and measuring the resource. Consequently, the exclusive property rights which are the basis of an exchange economy are difficult to establish and enforce.

Economies of large scale – Scale economies are evident in water storage, conveyance, and distribution. Therefore, water supply often provides the preconditions for a classic natural monopoly and, hence, water is generally supplied publicly or under regulation.

Variability in supply – Water supply is variable in time, space, and quality. The annual cycle of precipitation and streamflows prompts storage reservoirs to smooth out supplies. At the extremes of the probability distributions of availability, the unlikely event yields problems (floods, drought) which may be most economically solved when undertaken by public entities. Flood mitigation, for example, typically has public good characteristics.

Solvent properties – Plentiful supply and solvent properties create a capacity for assimilating and absorbing wastes and pollutants. Managing the assimilative capacity of the hydrologic system is, in essence, the allocation of a collective good, one that exhibits non-rivalry in consumption. It is this characteristic of water which requires the introduction of quality as well as quantity in the definition of use.

Sequential use – A given river may be tapped by many and varied entities as it flows from upper watershed to eventual destination in sea or sump. Only rarely is water fully consumed by any particular user. The "return flows" from upstream users may be reduced in quantity and degraded in quality, creating many problems for subsequent downstream interests, problems which require complex allocative institutions for solution.

Complementarity of outputs – Closely related to the previous point is the fact that some water may be used for more than one purpose. A reservoir can store water for flood control, irrigation, power generation, municipal demands, and recreation. Private ownership may capture only a part of these complementarities.

Bulkiness – Water is a "bulky" commodity, in that value per unit weight tends to be relatively low. Therefore, costs of transportation and storage tend to be high relative to economic value at the point of use, and the extensive transportation network developed to transport more valuable liquids (e.g. petroleum) is found only to a limited extent for water. This characteristic, combined with the relative costliness of enforcement of property institutions noted above, yields situations where the optimal property structure is the "commons" or open access.

Conflicting cultural and social values – Even where economic productivity might be best served by market allocations, alternative goals may oppose the result dictated by pure willingness to pay. Boulding (1980, p. 302) notes that "the sacredness of water as a symbol of ritual purity exempts it in some degree from the dirty rationality of the market". Market-induced shifts of water to energy or household uses which would alter flows or dry up streams are judged on the degree to which the natural environment or the existing social structure (i.e. the family farm) are affected. For such reasons, some cultures proscribe water allocation by market forces.

2.3. Public intervention in water resource allocation

2.3.1. The rationale for public intervention

Where markets are thin or absent, or where the demands or supplies revealed to markets capture only a portion of the full social costs of benefits, or when the commodity (water) in some role has public good characteristics, public intervention may allocate resources more efficiently. Public intervention may take a variety of forms: *regulations* (to provide for regularity of water use and to protect a given function of water – for example, recreation – against present and future competing demands; *public investment* in structures to protect against damages from flooding (a public "bad") or to provide infrastructure (for example, navigable water courses); or *public ownership and operation* to produce services jointly produced with other water related outputs (for example, hydroelectric power or municipal water supply). Collective action of these forms appears in a wide variety of combinations to serve a wide variety of objectives.

A number of facets of this issue, and the complexity involved, can be easily illustrated. Averting flooding through constructing a flood control dam yields a public good – when one downstream resident is protected from flooding, all downstream property owners are automatically protected. The provision of the dam may be socially worthwhile in that the social benefits may exceed the costs of building and maintaining the dam, but the private sector would fail to undertake the provision of flood control because of the difficulty of recovering costs from downstream beneficiaries. Similarly, some of the other "outputs" of water resource development may have public good or externality characteristics. These may be improved boating and picnic facilities created by the reservoir behind a dam, or beneficial side effects of a more reliable river channel or hydroelectric power potential created in constructing a flood control dam.

When goods involving these spillover effects are present, the efficient resolution often involves production by the public sector. Even in those cases where production is left in the private sector, public action may be necessary either to ensure the socially optimum amount of production or to correct for undesirable inefficiencies.

Thus, if the social benefit of these non-marketable services exceeds the cost of providing them, which it often does, and if the development of the river by a private firm precludes the development of these other purposes, which it often does, then private development of the stream denies society the benefit of these worthwhile yet external or spillover benefits. Multipurpose development by a government agency will permit society to enjoy the benefits of those products.

The converse of this may exist if private development imposes significant spillover costs. This is the case with proposals to construct hydroelectric or flood control dams which would flood out sites valuable for wilderness experience, scenic beauty, and other environmental values. In such a case, collective action may be required to keep a private project from being undertaken. It should be noted that this same conclusion would hold if the "developer" were a public agency rather than a private firm.

However, as Castle (1978) and Wolfe (1979) contend, government interventions may also "fail", so that combinations of market and non-market resource allocation mechanisms may yield the most appropriate solution in an imperfect world.

Finally, we can agree with Kelso (1967) who observes that while "water is different", the general public perception ascribes pecularities to water that go far beyond any idiosyncracies that can be objectively identified. Water policies and institutions are often out of touch with the realities of a world in which water is increasingly scarce. Even though water has special attributes, its allocation is an economic problem, and policies and institutions for its management should be designed to achieve economically efficient and equitable allocation.

2.3.2. *The nature of public intervention in the water sector*

Water management strategies may be distinguished according to several basic characteristics [White (1971)]. One characteristic concerns whether the water allocation decision is made by *public* or *private* sector decision-makers. Second, the project or program may be single-purpose or multiple-purpose. Third, the means employed may be distinguished as to whether one or more *techniques* or means are considered in providing project or program outputs. *Structural* or engineering approaches were the main forces of early policy, but *non-structural* or institutional means for solving water problems are receiving increasing attention. Finally, strategies may be judged according to a *single* criterion, such as economic efficiency, or *multiple objectives*, which may include the distribution of income or other social goals.

Federal intervention in the development and management of water resources in the United States dates from 1802, at which time the Corps of Engineers of the

U.S. Army was established. From the first Corps appropriation of $75 000 in 1824 – "for the removal of snags, sawyers, planters and other impediment of that nature" from the Ohio and Mississippi Rivers – public intervention, almost exclusively by the Federal government, has grown to enormous proportions. In the provision of irrigation water in the west, however, Federal legislation has also shaped the nature of ownership rights and market trading of water. Below, we briefly describe the nature and history of the public intervention in water resources in the United States. [See also Holmes (1972, 1979).]

2.3.3. Flood control

Although protection against flooding was one of the most recent water-related activities of government, it has firm economic rationale. A swollen watercourse has "public bad" characteristics – when flooding occurs, no downstream property owner or watercourse use is immune from damage. Conversely, an investment designed to reduce flooding, for example, a dam and reservoir, will automatically reduce damages from flooding for all downstream users. The Federal government has constructed numerous control reservoirs and dams, as well as undertaking river bed straightening and deepening and levee and revetment construction in areas subject to inundation.

Most Federal flood control expenditures prior to 1936 were administered by the Mississippi River Commission, mainly in response to the disastrous flood of 1927. The Congress in 1936 for the first time assumed nationwide responsibility for flood control, an activity which until then had been viewed as a local government responsibility (except for the Mississippi River). "Flood control on navigable waters or their tributaries is a proper activity of the Federal government." While flood control absorbed a relatively small share of Federal water resource expenditures in the period prior to 1945, during the post-war period, the flood control program became the major peacetime function of the Corps of Engineers. The Department of Agriculture, through its Soil Conservation Service, had the mission of reducing agricultural flood damages upstream of the large Corps installations on the major rivers. About one-half of total damages were agricultural. The expenditures of the Agriculture Department consisted of comprehensive soil conservation and land treatment activities and small storage projects in agricultural watersheds.

2.3.4. Navigation

Public investments in the inland waterway system have the same economic rationale as public infrastructure investments in transportation in general – the opening of undeveloped regions, the enabling of trade and communication among regions, and the provision of capital intensive right-of-way (with public good aspects). Water investments, historically, had a further purpose of stimulating a traffic mode which could effectively counter the monopoly power and preemptive practices of the railroads. The Federal navigation program has been the domain of the Corps of Engineers, and has been focused historically on the Great Lakes ports and inland waterway system. In the past two decades, however, activities have concentrated on the Mississippi and Ohio Rivers, ocean harbors, and coastal rivers.

The inland waterway program consists of a wide variety of project types – dams designed to regulate flows to navigable depths, dredging and straightening watercourses to permit barge transportation, the construction of canals where no natural watercourse exists, the construction of Great Lake and ocean port facilities, and the maintenance of all of these.

2.3.5. Hydroelectric power

Public production of electric power is largely a post Second World War phenomenon. With few exceptions, electric power generation is a secondary purpose of projects the primary function of which is to provide flood control, irrigation or navigation services. The hydroelectric generation function has been justified as an economical by-product of irrigation, flood control and navigation projects. The Corps of Engineers is responsible for only about 20 percent of the public hydroelectric capacity in the United States; the primary construction agencies are the Tennessee Valley Authority (over one half of the capacity) and the Bureau of Reclamation which has provided hydroelectric generation as part of a number of large irrigation projects in the western states. While the Corps of Engineers and Bureau of Reclamation have been responsible for project construction, the power is marketed through special agencies in the Department of Interior with a cost recovery mandate.

2.3.6. Irrigation

The irrigation program began with the Reclamation Act of 1902, which authorized the Bureau of Reclamation of the U.S. Department of Interior to build irrigation projects. The reclamation program is confined to the 17 western states (plus Alaska), and is financed by sales of public lands, beneficiaries of projects (which are required to pay some share of the costs), the sale of electricity and general appropriations. While the Bureau of Reclamation is responsible for construction of the projects and the arrangement for reimbursement, management and maintenance is turned over to user-managed irrigation districts.

The economic rationale for irrigation investments is one of the most tenuous of the Federal water resource activities. Three reasons have been suggested: (1) the infrastructure, regional development rationale, (2) the need for eminent domain rights in order to secure water rights and land rights for project construction, and (3) the massive initial capital requirement that creates a barrier to private or local provision. However significant these may have been in the west in the early part of the century, they are of questionable import now.

2.3.7. *Other water resources purposes*

In addition to the interventions described above, a range of other public activities involving the provision or use of water have been undertaken by the public sector. Here these will simply be mentioned.

Water pollution control. The Federal role in water pollution control was trivial before 1960, and modest until 1966. During the early 1960s, these activities were under the supervision of the Department of Health, Education, and Welfare (HEW) and included: data collection and dissemination, research, administration of pollution control grants to state and local governments and industry, and enforcement procedure) of the water pollution control act. The economic rationale for this intervention is clear: reducing or preventing spillover costs on downstream water users from the discharge of effluents.

During the decade of the 1960s, the organizational arrangements for pursuing water pollution control changed substantially, culminating in the creation of the Environmental Protection Agency (EPA) in 1970. The functions of the pollution control offices (and the appropriations granted them to support these functions) were expanded to include: extended enforcement powers, establishment of water quality standards for all watercourses (including the setting of criteria and a plan to implement the criteria), and (after 1970) the setting of effluent standards and the enforcement of the standards on both municipalities and states. Throughout the entire period, the strategy of the Federal government was basically two-pronged: the setting and enforcement of regulations (standards) and the provision of subsidies to accelerate pollution control activities.

Municipal and industrial water supply. The provision of water to municipalities and industrial users has been a long-standing by-product of the Reclamation program. Beginning in the 1960s, however, such deliveries and contracts became more important than in earlier periods, but remained but a small fraction of irrigation water deliveries.

Recreation. As with the water supply function, the Federal provision of recreation services has also grown, again largely as an economic by-product of activities whose basic purpose was flood control, navigation, or irrigation. The various agencies have accepted this function, and implemented it by the construction of parks and recreation grounds adjacent to reservoir facilities and the provision of access to and the regulation of water based recreation activities.

2.4. The objectives of public intervention in water resource allocation

The prime requisite for evaluating public interventions in the water resources area is an explicit statement of the objective toward which the resource development decision is focused. The benefits attributable to use of a resource have meaning only in relation to the objective, and are measured as the contribution of the resource to the objective function [Marglin (1962)]. Discerning the objectives of public interventions in the water resources area is difficult as the principal public sector decision makers often fail to articulate any clear purpose for their decisions. Nevertheless, the statements and actions of policymakers do seem to point rather systematically to the interaction of two objectives which guide public interventions: (1) economic efficiency and (2) regional economic aid or income redistribution.

From the very inception of Federal government activity in both the development of navigation facilities and flood measures, some emphasis has been placed on the degree of economic efficiency of the projects to be constructed. While tangible evidence of such concern is found earlier, the Congress in 1936 further reaffirmed and clarified this position by requiring that, for such projects to be authorized, benefits must exceed costs, "to whomsoever they may accrue". Since that time all water resource projects have been evaluated by the evolving methods of benefit–cost analysis.

While concern with economic efficiency is of long standing in the history of water resource development in the United States, other criteria have also been explicitly recognized, in particular, income distribution and regional development. The concern with the multiple objectives to be served by public water resource developments is reflected in both the writings of scholars in this area and in official government documents.

In an early statement (1952), the Bureau of the Budget's *Circular A-47*, discussing the criteria to be applied by the executive office in the review of project reports, placed great emphasis upon economic efficiency in defining concepts to be included as benefits and costs. Also, however, "the efficiency of the program or project in meeting regional...needs" is stated as a further criterion. The "Green Book" [Federal Interagency Committee on Water Resources (1958)], while again heavily emphasizing the necessity of total annual benefits exceeding estimated annual costs, explicitly noted the importance of regional development as a public water policy objective.

This growing recognition on the multi-dimensional nature of the social welfare function in planning for water resources was extended in Senate Document 97 [U.S. Congress (1962)] and was formalized in the Water Resources Council's *Principles and Standards*, in 1973. The 1973 *Principles* established four accounts on which evaluation was to be based – national economic development (economic efficiency), regional development, environmental quality and social well-being. These categories were maintained in the documents of the Water Resources Council's (1979) revisions and extension, while procedures for measuring beneficial and adverse impacts were refined.

The U.S. Water Resources Council's (1983) *Principles and Guidelines* retained the same four-account classification with some minor changes in nomenclature and procedure. This document returns the emphasis to the national economic development objective while requiring plans to be consistent with environmental protection.

In sum, then, the focus on economic efficiency – the existence of project benefits in excess of costs – in water resources has been fundamental and persistent. However, a basic and growing tension between this efficiency goal and other objectives – largely, regional development or income redistribution – exists. [See Eisel et al. (1982) and Castle et al. (1981) for more detailed discussions of the evolution of Federal evaluation procedures.]

3. Benefit–cost analysis for water resources systems

3.1. Conceptual basis

The prevailing technique for evaluating public investments and policies in the water resources area is benefit–cost analysis. This approach assumes that economic efficiency is the relevant objective for public water resources interventions. Procedures for estimating the benefits and costs of a non-marketed commodity such as water can be interpreted as efforts to simulate hypothetical market outcomes. The basic concept of "benefit" underlying such estimation is *the amount a rational and informed user of a publicly supplied good would be willing to pay for it*. Costs represent the forgone value of goods and services displaced by a project. [See one of the general texts on cost–benefit analysis, e.g. Pearce and Nash (1981), Mishan (1976) or Gittinger (1982) for more general treatments of the subject.]

Willingness to pay, which reflects the user's willingness to forego other consumption, is in turn, formally represented by a demand curve relating the quantity of a good taken at a series of alternative prices. [The producer's demand for an input is given by the marginal value product (MVP) for that input.] The value of additional units decreases as the quantity consumed increases. The negative slope of the demand curve follows from the principles of diminishing marginal utility of consumers and diminishing marginal product for producers. The measurement of willingness to pay should be designed to be consistent with market prices.

Krutilla and Eckstein (1958) presented a conceptual framework for analyzing mutli-purpose river basin investments. Marglin (1962) formalized the model, and extended it to more complex cases where demands are interdependent and budget constraints apply. A net benefit criterion function, representing the present value of the streams of future benefits and costs, is maximized. [See Herfindahl and Kneese (1974) for a succinct summary.] Marglin's summary provided the basis for developing interdisciplinary computer simulations models [Maass et al. (1962), Hufschmidt and Fiering (1966)], which played an influential role in the subsequent development of the water resource planning literature.

While computer simulations may employ the model of optimal resource allocation, various simplified formulas are employed to measure project worth at the field level. These include the net present value, the benefit cost ratio and the internal rate of return. Any text on cost benefit analysis describes their computation, use and limitations [James and Lee (1971), Gittinger (1982)].

3.2. Problems in measuring the economic impacts of water resources interventions: Conceptual issues

There are a number of conceptual issues relating to the general question of measuring the impacts of water resource interventions, to which we now turn.

3.2.1. "With or without" principle

This rule asserts that benefits and costs are to be measured as increments which would occur *with* the project or program as compared to without. Adherence to the rule assures that measured benefits (or costs) are solely due to the program or project, rather than measures of changes between *before* the project as compared to *after*, some of which would have occurred autonomously even in the absence of the program.

3.2.2. The accounting stance

In the theoretical construct of the market system, a private accounting stance is presumed. Individuals are motivated to act in accordance with gains and losses as each perceives them, and pursuit of private objectives (such as maximizing utility or profits) is assumed to occur independently of gains and losses occurring elsewhere in the system. When the responsibility for an allocation decision rests with a public agency, an alternative criterion may be appropriate. In the water resources literature, two major alternatives to the private perspective are found (i.e. alternative "objective functions" or "accounting stances"). These reflect the viewpoints, respectively, of regional planning authorities (river basin or state) and the Federal government [Howe (1971, ch. 2)].

Regional and national accounting stances differ from private financial analysis in that social rather than private benefits and costs are incorporated into the analysis. Ideally, the national accounting stance should attempt to utilize social opportunity costs and values for all inputs and outputs, whether they are correctly or incorrectly priced by the market mechanism, or not priced at all. All externalities should be identified and incorporated into the measures.

3.2.3. The equimarginal principle

The marginal benefit represents the contribution of an incremental unit of good or factor to a specific objective function and is defined by the first derivative of the total benefit function. As was shown above, it is the net marginal benefit function which is of primary importance for purposes of efficiency analysis in water resource development and allocation. For the development case, economic efficiency requires that development be undertaken to the point of equality between the marginal value of the output and its marginal cost. For the reallocation decision (i.e. the allocation of constrained water supplies among competing uses), economic efficiency is achieved when net marginal benefits per unit of water are equal for all uses. This latter proposition is familiarly known as the equimarginal principle.

3.2.4. Long-run versus short-run value

A fourth conceptual distinction is that between short- and long-run value. This distinction is related to the degree of fixity of certain resources and is especially important where commodities are used for further production (i.e. intermediate as opposed to final goods), as is typical with respect to water.

The rational producer's willingness to pay for an increment to water supply is equivalent to the increase in the net value of output attributable to the added water. The distinction between short-run and long-run value is that in the short run, where some inputs associated with water use are fixed, estimates of increases in the net value of output can appropriately ignore the sunk costs of the fixed resources. However, in the long run, all costs must be covered.

3.2.5. Physical interdependence and economic impacts

The above discussion points to a major problem which increases the difficulty of evaluating the benefits and costs of using water. A specific water use cannot, in most cases, be viewed in isolation from potential alternative utilizations. The typical river basin will contain several alternative uses for water and one use may affect others through any or all of the quantity, quality, time, and location dimensions. The benefits from a particular increment of water supply in a given river system is the sum of the value of the marginal product in the initial use and the value of the return flow in all subsequent uses. In a system context, the sum is net of the positive and negative effects which are engendered elsewhere or subsequently in the system [Hartman and Seastone (1970), Butcher, Crosby and Whittlesey (1972)].

3.2.6. Appropriate measure of use

Formally, the valuation problem posed by physical interdependencies in water use is that of specifying the unit of measure of the variable representing quantity of water. In certain situations (e.g. complementary products such as recreation), evaluation of water resource development decisions may not require a measure of value per unit of water "used". This is true so long as a use is not competitive with another.

Another problem is posed by instream utilization. While navigation, recreation, power generation, and waste load assimilation do not withdraw or consume water in the usual hydrological sense of these words (evaporation and seepage aside), instream uses clearly can foreclose other economic uses at a particular location and at later times. The "with and without" principle will provide guidance in such instances.

For cases involving withdrawal use, some unit measure of use is clearly required for the evaluation of alternative uses. The choice of the appropriate measures of use is typically between the withdrawal versus the depletion (consumption) concepts. D'Arge (1970) contended, for example, that the selection of the appropriate variable depends on the interdependencies existing among users and on the availability of benefit estimates. He concluded that consumption is the relevant variable for public planning purposes.

Most economists seem to prefer to measure use in terms of withdrawal, since that is what the individual private user must pay for. Moreover, conventional measures of consumption, in terms of evaporation may be misleading for economic analysis from a social perspective. As examples, return flows may so degrade in quality as to be unusable or return flows percolating back to the groundwater table in deep aquifer situations may not be available for reuse in any reasonable planning horizon.

3.2.7. Economic benefits related to location, quality, and time

Site productivity refers to the economic value of water used at a particular geographic location for a specific class of use. The costs of transporting this bulky commodity are such that the derived economic value of water in the stream will be considerably less than at the point of use. The factors which influence the physical productivity of water at a particular location for each type of use also affect the economic value. Examples of these factors include soil and climatic characteristics affecting the physical productivity of irrigation water or aesthetic characteristics of a particular site which influence the value of water for recreation. The productivity of water is also dependent upon the degree of investment in other resources used in conjunction with water, such as the height to which a power dam is constructed or the investment in efficient water application systems in irrigation.

Temporal variability in demand can significantly affect benefit estimates. The variation may extend from the very short run to the long run. The most important case stems from seasonal variation such as shifts found in demand from agriculture, navigation, recreation, and waste load assimilation. Also, secular trends in population, income, and technology have a long-run impact on the demand for water.

Water must frequently undergo some form of processing (filtration, chlorination, pressurization, etc.) prior to use. Thus, there will be differences in willingness to pay for the raw (unprocessed) water as compared to the benefits of water of suitable quality for a specific use.

To sum up, specifying strictly commensurable shadow prices for alternative uses of water requires that benefits per unit of water be conceptually equivalent in terms of time, location, and quality. [Flinn and Guise (1970) and Howitt et al. (1982) present sophisticated modeling efforts which incorporate these distinctions.]

3.3. Techniques for determining the direct economic impacts of water resources interventions

Five broad approaches for measuring the benefits of water resources interventions may be identified [Young et al. (1972), Gray and Young (1984)]. These include: (a) observation of transactions relating to water, (b) derivation of value from a statistical demand function, (c) residual imputation and variations, (d) alternative cost valuation and (e) user surveys. (Where certain costs of water development projects are not correctly reflected in market prices, these techniques are also applicable to the measurement of such costs.)

3.3.1. Market transactions relating to water

Because of the physical, economic, and institutional characteristics of water, market transactions for water are rare. However, they do exist and in such cases the observed price must be carefully interpreted. The least complex has been termed the "irrigation water rental market" [Anderson (1961)]. The owner maintains the title to the perpetual annual stream of water supplies but sells the right to receive the water for a specified period of time. The observed prices in rental markets are based on private, short-run demands and may be of limited utility in evaluating long-term public investment or reallocation decisions.

Transactions in permanent water rights are not common, largely because of institutional constraints designed to avoid third party effects [Ditwiler (1975)]. Observed transaction prices of transfers between similar uses are conceptually correct measures of the long-term *private* value of the resource in that purpose. However, interpretation of these derived values must be done with care if public shadow prices are needed. Also, the appropriate price is that for the right to a perpetual series of annual flows, and not the price of a unit volume of water. In order to derive the value of a unit volume, an appropriate capitalization formula, with the proper interest rate, must be applied to the price of the right. Will the private sector exhibit the same rate of discount, risk aversion, time horizon or price expectations as would be selected by the public analyst? Gardner and Miller (1983) have illustrated this approach with data from the Northern Colorado Water Conservancy District, while Brown et al. (1982) have studied markets for water rights in New Mexico. The cyclical price variations observed by Gardner and Miller are consistent with the hypothesis that the market price of water rights can be affected by the same imperfect forecasts of inflation or urban growth rates as are markets for precious metals, real estate, or common stocks.

The value of water rights has also been estimated indirectly where the right is transferred as a part of a real property transfer. Statistical regression analysis applied to a sample of such transactions characterized by variation in water supply per unit of land permits inferences to be drawn as to the capitalized value of the water right. Freeman (1979, ch. 6) presents a detailed review of the problem of employing property values to study the benefits of non-marketed goods and services, particularly with respect to environmental quality. (See also Chapter 15 of this Handbook.)

A second type of observed transaction in water is that in which water supplies in withdrawal uses are sold under an "administered" price system. In this case, the public agency or utility which supplies water may sell it at a specified price through a metered system. The consumer is free to adjust consumption to reflect the marginal valuation of water use at the specified price. Statistical analysis of cross-section or time-series data pertaining to the consequent relationship between consumption and price can be used for inferring the value of water to the final user [Howe (1982), C.E. Young et al. (1982)].

3.3.2. *Benefit estimates from econometric production functions*

The classical approach to estimating values of non-marketed commodities is to estimate the demand function for the good in question. Water is often an intermediate good, in which case the demand function is the marginal value product function, the first derivative of the production function in value terms. Use of econometric production function estimates is most common in valuing water in irrigation use, where numerous field experiments have studied crop response to water application and other factors [for example, Hexem and Heady (1978)]. The general approach is to derive a schedule representing the short-run value of the marginal product under the experimental conditions.

Cobb–Douglas type functions may be fitted to farm account data with irrigation water as an explicit variable have been employed in estimating long-run marginal value productivity. A major problem in such cases is obtaining an accurate measure of water applied. Extrapolation, of course, must be done with caution. Moreover, the derived value may not be suitable for social cost–benefit analysis if commodity market interventions or unemployment are present. Ruttan (1965) and Beattie et al. (1971) have applied the approach to aggregate irrigation and production data.

In industries other than irrigated agriculture, a scarcity of the data necessary to estimate demand or production relationships, together with the fact that water accounts for a very small portion of production costs, has generally forced analysts to turn to alternative estimating procedures.

3.3.3. *Residual valuation approaches*

Residual imputation achieves the task of shadow pricing by allocating the total value of output among each of the resources used in a single productive process when water is used as an intermediate good. If appropriate prices can be assigned to all inputs but one, then the residual of the total value of product is imputed to the remaining resource [Heady (1952)].

The residual imputation technique is based upon two major postulates: (1) the market prices of all resources, except the one to be valued, are equal to the returns at the margin (value of the marginal product), and (2) the total value of output can be divided into shares such that each resource is paid according to its marginal productivity and the total value of output is completely exhausted (Euler's Theorem). Consider a simple example where three factors, capital, labor, and water, are used in the production of a single output Q. The problem is to impute a value to the water resource. By Euler's Theorem:

$$TVP_Q = VMP_L \cdot L + VMP_K \cdot K + VMP_W \cdot W \tag{1}$$

where TVP_Q is the total value of output Q, VMP_i represents the value marginal product of any resource, i, and L, K, and W refer, respectively, to quantities of labor, capital, and water employed. Substituting according to the first postulate and rearranging, we have

$$TVP_Q - P_L \cdot L - P_K \cdot K = VMP_W \cdot W. \tag{2}$$

Eq. (2) is solved for VMP_W to estimate $\hat{P}_W$, the desired shadow price of water.

The postulates cited previously are satisfied by production functions homogeneous of the first degree and optimizing producers in competitive markets. The Cobb–Douglas function, which implies constant returns to scale, satisfies Euler's Theorem and has been used in empirical estimation of marginal value products.

Residual imputation is subject to limitations which should be recognized by the user. First, if input variables are omitted, inadvertently or otherwise, the residual will be overstated. Second, distortions of either input or output prices will lead to a distorted residual estimate. Lastly, this procedure is most applicable to estimating the value of water in production processes (such as irrigated crop production), where the water resource is a substantial contributor to total product. In industrial uses, where the contribution of water rarely represents more than 1 or 2 percent of total value of product, the difficulty in properly shadow pricing the other factors, particularly capital, management, and risk-bearing, leads to highly uncertain estimates of a residual value of water.

Mathematical programming procedures can be employed to derive imputations of the value of water which are theoretically similar. Burt (1964) pioneered this approach with application to irrigation water, deriving a long-run net benefit function from parametric variation of a water supply constraint in a linear programming (LP) model of a California agricultural region. Depending on the definition of the objective function, long-run or short-run value estimates are obtained. Numerous others have used LP models to impute short-run values to irrigation water, in which case the residual is the return to land, management and fixed investments, in addition to water.

The "Change in Net Income" method (hereafter abbreviated to CINI) is related to the residual valuation approach. This model defines the increment in net producer income associated with adding water to a production process as willingness to pay for the incremental water. The approach is that adopted for valuing irrigation water benefits by the U.S. Water Resources Council (1979). Assign X_j to represent production inputs and Y_i refer to products, and let the subscripts 0 and 1 attached to the input and output variables refer, respectively, to values *without* and *with* and investment or program adding to water supply. The water resource is designated X_1. Assuming that the factor prices (P_{x_i}) and product prices (P_{y_i}) are unaffected by the investment, the change in net income ΔZ associated with a discrete addition to water supply per unit of time is:

$$\Delta Z = Z_1 - Z_0 = \left(\sum_{i=1}^{m} Y_{1i} P_{y_i} - \sum_{j=2}^{n} x_{1j} P_{x_j} \right) - \left(\sum_{i=1}^{m} Y_{0i} P_{y_i} - \sum_{j=2}^{n} X_{0j} P_{x_j} \right). \qquad (3)$$

The second term in (3), in effect, represents the annual net returns to the fixed land resources in the "without" project situation.

The unit value of water may be obtained by dividing the expression in equation (3) by the incremental quantity of water (i.e. ΔX_1).

The CINI approach requires the same assumptions of the residual imputation procedure, namely, that resources be optimally allocated, that factor and product prices correctly reflect social values, and that all inputs be properly represented in the calculations.

A number of studies have attempted to measure the value of water from a regional perspective employing regional inter-industry models. Such studies typically employ a concept of "value added", or more generally, of income of primary resources per unit of water withdrawal as a criterion for allocating the resource.

What we term here as the "value-added approach" has certain key similarities to the residual approach described above, but it also has important differences. Although its practitioners have presented it as an appropriate method for valuing an unpriced resource such as water, this interpretation can be accepted only under very limited conditions. The important difference between the value added and the correct residual approach lies in the definition of value added. Since value added is generally an aggregation of the basic inputs to production, the residual in this case includes not only the contribution of water to the value of output, but the contribution of all primary resources. Attributing the value added to water implicitly assigns a zero shadow price to the other primary resources and thus ignores the fact that resources other than water are scarce. Assigning zero opportunity cost to other primary resources by implicit assumption is questionable, and tends to result in water value estimates which greatly overstate the true contribution of water to net regional output. The value-added imputation process can lead to conceptually correct results only if (1) extreme care is taken to disaggregate value added so that the contribution of all other primary resources is empirically identified and deducted from value added or (2) if the assumption that the opportunity costs of the other primary factors is zero is verified. A number of well-known studies by regional economists have used this method, and their results are subject to this critique [i.e. Wollman (1963), Lofting and McGauhey (1967), d'Arge (1970), and Bradley and Gander (1968)].

3.3.4. *Alternative cost*

The fourth major technique of value estimation discussed here is based on the concept of "alternative cost". Alternative, in the alternative cost context, refers to a substantively different means of accomplishing the same project purpose. Willingness to pay is limited to the cost of the most likely economically feasible alternative. The definition is deceptively simple because there are a number of possible alternatives, including private alternatives to public projects, public alternatives to each component of dual purpose projects, and so on. [See Herfindahl and Kneese (1974, pp. 267–270), for a more detailed exposition.] The technique is applicable to cases in which a private alternative (e.g. a railroad system for commercial transport) to a public development (e.g. navigation for the same purpose) exists. Maximum willingness to pay is determined by the cost of the least expensive alternative.

The approach is attractive since in many cases estimation of a demand schedule is very difficult, if not impossible. However, complexities arise in the situation in which neither alternative need be built to a fixed scale. Then the demand schedule must be estimated for the output range between the private level of output and the public level of output (assuming demand is not totally inelastic) of the private alternative represents the upper limit of willingness to pay (benefits) for the public alternative.

The primary advantage of the alternative cost method is that, for cases in which demands are difficult to obtain, estimation of maximum willingness to pay can be accomplished without estimating demand functions. In those situations where the output of each of two alternatives is water, as in the case of private development of groundwater for irrigation versus public supply, the least cost alternative can represent a legitimate estimate of the social value of water. In some situations, e.g. transportation, power production, and waste treatment, estimation of direct benefits to water, as contrasted with total project benefits, is a two-step process. First, the alternative cost of accomplishing a given purpose must be estimated. Second, a benefit per unit of water must be imputed, usually by deducting from the alternative cost the associated costs (an application of the residual technique described above).

3.3.5. User surveys

The final category concerns methods for determining the demand for water when no exchange transaction or diversions for production occurs, that is, when the "use" activity involves neither consumption nor diversion. In such cases, usually associated with recreation and aesthetic enjoyment of water in natural surroundings, water has a public or collective good character. Here, analysts have come to rely on user surveys to derive estimates of the value of the recreation experience, and more particularly, of the value of the contribution of environmental resources, such as water, to that experience [Knetsch (1974)]. Two general lines of approach can be identified – the expenditure function approach, and the income compensation approach.

The expenditure function approach relies on market-generated price and quantity data where the quantity of a non-rival good is an argument in the demand for some private good. Under certain conditions regarding the demand relationships, an empirical estimate of the benefit for the non-rival commodity can be derived. The well-known Clawson–Knetsch travel cost method and land-value approaches to valuing amenities are examples.

The income compensation function approach derives from the Hicksian model of monetary equivalent measures of welfare impact. Willingness to pay is defined as the area under the Hicksian compensated demand curve. Since the indifference surfaces of the theoretical model are not directly observable, various approaches to estimation have been developed. These mainly depend on a "direct asking" (contingent valuation) approach to estimating changes in economic surplus. Freeman (Chapter 6 of this Handbook) develops these issues in detail, and we do not treat them further.

3.4. Social cost measurement

For the most part, water project appraisal proceeds on the assumption that the relevant markets reasonably accurately reflect the costs of factor services and intermediate goods employed by the public sector. In smoothly functioning markets, wage rates and the prices of material and equipment adequately measure the opportunity cost of resources. Some possible exceptions to these presumptions are noted here.

3.4.1. Underutilized resources

The existence of unemployed and underutilized resources must be recognized in any attempt to estimate the true social cost of a water resource project. The opportunity cost of underutilized resources is less than the market price, since little forgone production occurs if such resources are utilized.

Early writers [Eckstein (1958), McKean (1958)] acknowledged the problem, particularly with regard to labor, but were skeptical of attempts to forecast unemployment over the long interval between project planning and construction or the even longer period of the operating project's lifetime. Haveman and Krutilla (1968) initiated efforts to develop empirical measures of the opportunity costs of underutilized labor. A response function was formulated which relates the probability of drawing from a pool of idle resources to the unemployment rate in that pool. Adopting an *a priori* hypothesized form of the response function, they concluded that the true social cost of projects were some 5–30 percent less than monetary costs (based on employment data from 1957 to 1964). Significant regional variation was found in the appropriate adjustment factor. They and others note, however, that for appropriate measurement of the social cost of labor, it is not sufficient to measure the "before" and "after" labor force status, but that "with" and "without" conditions must be compared. The latter is particularly difficult to achieve, particularly given the prospect for labor migration from other regions.

Federal evaluation procedures regarding underutilized resources have varied rather widely. Senate Document 97 [U.S. Congress (1962)] encouraged accounting for project construction period and operating period unemployment, and permitted recognition of indirect and induced employment effects. The U.S. Water Resources Council's (1973) Principles and Standards limited consideration of underutilization to only the construction or installation period, a practice continued in the (1979) and (1983) revisions.

We close this topic by calling attention to a rather different approach. Johnson and Layard (1982) adopt a general equilibrium framework, and show under plausible assumptions that the social opportunity cost of unemployed labor can exceed the wage rate.

3.4.2. Opportunity costs of non-marketed resources

Market prices may be biased or absent in the case of lands used for project sites. Recreational benefit forgone from reservoir construction is an example. If the site is publicly owned, no budget outlay for its purchase is required, but non-marketed alternative (e.g. recreational) uses may be sacrificed. (If the land is purchased, the higher private discount rate, an aversion to risk and differing price expectations on the part of private land market participants may imply an assessed value which diverges from that derived from evaluations from a social accounting stance. Particularly in instances where an irrigation project inundates some farm lands to develop other farm lands, the opportunity costs of the site should be assessed with the same discount rate, commodity prices and production costs as are the benefits of the development.)

Finally, the opportunity cost of the water itself must not be ignored. As water economies mature and the resource becomes increasingly scarce, the potential forgone values will rise. Since relative economic values and institutional arrangements both act to protect household and industrial demands, the problem arises mostly with instream uses. Hydropower benefits forgone can be very high when irrigation water is diverted high in a river basin [Whittlesey and Gibbs (1978)]. Recreational uses requiring instream flows are only beginning to be protected by legal rights [Daubert and Young (1981)].

3.5. Other benefit–cost analysis issues

We treat briefly below a number of long standing economic issues which have pervaded evaluations of public interventions in the water field.

3.5.1. The discount rate

> Around the mysteries of finance
> We must perform a ritual dance
> Because the long-term interest rate
> Determines any project's fate:
> At two percent the case is clear,
> At three, some sneaking doubts appear,
> At four, it draws its final breath
> While five percent is certain death.
>
> Kenneth Boulding (1964)

A long-standing issue in applying benefit–cost analysis in the water resources area concerns the rate of interest to be used in the discounting of future streams of benefits and costs. Until 1969, official procedures manuals stipulated that "average long-term interest rates that will prevail over the life of a project are considered the proper basis for discounting future benefits and costs". The long-term government bond rate was taken as the measure.

This rate was rationalized as follows. First, it is claimed that the rate of interest conceptually appropriate for use by the Federal government is the social cost of capital, i.e. "... the risk-free return expected to be realized on capital invested in alternative uses". Second, because the government can borrow funds at the going long-term government bond rate, it is claimed that this rate is an accurate estimate of the social cost of capital. Critics of this procedure have argued that because of the difference between lender's and borrower's risk, the rate on long-term government issues is less than the social opportunity cost of capital [Eckstein (1958)]. Also, due to the effect of taxes, the actual bond rate may fall far short of the real opportunity cost of capital as determined by its pre-tax value in an alternative use. (Offsetting these biases in recent years has been the effect of inflationary expectations and heavy government borrowing in increasing the nominal cost of capital to the government.)

The appropriate conceptual basis for the discount rate to be used by the public sector has been long-debated in the economics literature. Depending on the perspective taken, a case can be made for any of the following concepts: (1) the social rate of time preference [in conjunction with a cut-off benefit–cost ratio to reflect opportunity costs (Marglin (1968), Eckstein (1958))]; (2) the opportunity cost of displaced private spending, [Haveman (1969), Harberger (1968)]; and (3) the before tax rate of return on corporate investments [Stockfisch (1982)].

In 1968, H.P. Caulfield, Director of the Water Resource Council, proposed a formula approach to determining the discount rate on federal water resources investments. After Congressional hearings, the Water Resources Council announced that the new formula was to be based on the *yield* rate for long-term government bonds. An immediate 4.625 percent rate was established (in comparison to the 2.5–3.5 percent rate in effect during the 1950s and 1960s) and the much higher yield rate was to be approached from this level by not more than one quarter percent per year. In fiscal year 1984, the rate used by the agencies, based upon this formula, had risen to 8.125 percent. This rate has come to closely approximate the real opportunity cost of displaced private spending which is preferred by many economists. However, the conceptual basis on which the water resources rate is based – the nominal cost to the Treasury of borrowing – is little different than that on which the pre-1970 rate was based.

The debate in economics on the appropriate concept and magnitude for discounting public investments has continued but there is still little concensus on the correct concept and size of the discount rate. Nevertheless, most economists would agree that the cost of Treasury borrowing is not an appropriate conceptual basis and that, in the analysis, the opportunity cost of alternative activities displaced by the public activity and its financing must be considered. A range of other issues pertinent to the choice of the discount rate have not been resolved, however. These include: (1) the appropriate consideration of future generations, (2) the inclusion of risk and uncertainty considerations in the discount rate, and (3) the relationship of the appropriate public discount rate to macroeconomic policy. These issues are discussed fully in Lind (1982).

3.5.2. Inflation and benefit–cost analysis

The worldwide experience with inflation in the seventies raised questions regarding the treatment of inflation when performing B/C analyses. Price and interest rate data used in such investigations often reflect inflationary expectations. The major conclusions of the literature on this subject [Howe (1971), Hanke et al. (1975)] is that consistency is required in the treatment of prices and interest rates. In other words, either real or nominal prices and interest rates must be considered in making projections. (The Water Resource Council procedures continue to violate this precept by appealing to nominal interest rates – the yield rate on long-term government bonds – while employing real prices in forecasting benefits and costs.) It is recommended that real values be employed since the forecasting of nominal price trends over the long lifetime of a water project is a formidable task. Hanke and Wentworth (1981) extend the analysis to cases in which relative prices are expected to change over the life of a project. (The writers would add a note of caution on projecting changing relative prices. Analysts who in the 1970s confidently adjusted relative prices of food or energy in response to perceived permanent scarcity scenarios for these commodities have seen supplies rebound and real prices in the 1980s fall to well below earlier forecasts.)

3.5.3. External and secondary effects

Effects internal to a water project are those which can be captured, priced, and sold by the decision-making or project entity (or which must be paid for, in the case of costs). External effects, then, are uncompensated side effects, and can be positive or negative.

External effects can be classified as technological or pecuniary. The former refer to changes in real production or consumption opportunities imposed on third parties, and generally involve some physical interaction among the parties. Because technological externalities are real and represent welfare changes, public project planning should take them into account. Pecuniary impacts (usually called "secondary" or "indirect" economic effects) are those reflected in changes in incomes or prices caused by shifts in supply or demand. Pecuniary externalities are likely to represent income distribution rather than allocative effects, and their inclusion would amount to double-counting.

The need to explicitly consider real effects on third parties in benefit–cost evaluations is clear and presents no serious conceptual difficulties. Whether positive or negative, such impacts can be measured, in principle, by methods of non-market valuation discussed above in Section 3.3. In practice, of course, serious difficulties in measuring external costs and benefits abound.

Pecuniary spillovers present more of a problem. The conditions under which pecuniary externalities are properly included in measures of water project benefits have been the subject of a long and controversial history. McKean (1958) and Eckstein (1958) remain the definitive analyses. The main problem has been a focus by planning agencies on secondary benefits (which are largely registered in the project locality), while secondary costs (which are likely to be spread across the national economy, and often represent the elusive potential returns to alternative public investments) are not given equal consideration.

Howe and Easter (1971, pp. 26–27) present a most accessible summary of the issues. They note that in a properly functioning economy with fully employed resources, a new investment project yields no net benefits beyond its own net income. Any expansion in related activities is offset by a fall in activity and profits elsewhere, while potential alternative investment projects could be expected to have similar indirect effects. However, with departures from the competitive model – including (a) the presence of long-term unemployment of resources, (b) immobility of resources, and/or (c) the existence of economies of large scale in related industries – real national secondary impacts may occur.

Two final remarks are in order. Even though the pecuniary (or secondary) impacts of water projects are likely to be balanced out elsewhere in the national economy, that is not to downplay their economic and political importance to affected regions [Kimball and Castle (1963)]. Much of the political motivation for public water projects represents an attempt to capture such regional effects, which in many cases are reflected in large gains in real property values. Second, much analytic effort has gone into forecasting regional income gains, often with the use of Leontief input–output models. The changes thus projected remain income transfers, and should not be labeled as "benefits" or treated as real income gains.

Kelso et al. (1973) and Bell et al. (1982) are empirical measurement efforts which avoid the possible pitfalls in projecting secondary impacts.

3.5.4. Risk and uncertainty

Risk and uncertainty attached to outcomes – either positive or negative outcomes – are generally viewed as undesirable from an individual point of view. Individuals are generally thought of as "risk averse". Production and consumption decisions made in the face of uncertainty are generally less effective than when knowledge is relatively certain [Dorfman (1962)]. Extrapolation of these points to project evaluation implies that risky investments – those in which benefit and cost streams are largely uncertain or risky – are less desirable than interventions with equivalent expected values of benefits and costs, but less risk or uncertainty. The implication is that risk and uncertainty is a cost, and this cost should be reflected in the evaluation of projects.

Existing evaluation procedures used by water investment agencies, as well as the academic literature, reflect this conclusion. A number of approaches to reflecting risk and uncertainty have been proposed, including: (a) conservative "rules of thumb", (b) sensitivity analysis, (c) probability analysis, and (d) decision-theoretic models. The "rules of thumb" are devices for penalizing riskier proposals, and include (1) limiting the period of analysis, (2) introducing direct and specific safety allowances, (3) the inclusion of a risk premium in a single discount rate (where uncertainty is unrelated to time), and (4) appraising benefits conservatively. No longer does the project appraisal literature [Mishan (1976)] or the U.S. Water Resources Council Guidelines (1983) recommend these approaches. They do, however, advocate sensitivity analysis, which is a reworking of the analysis for alternative values of the parameters which are thought likely to affect the feasibility determination.

The probabilistic approach generally relies on formal assignment of probabilities to uncertain outcomes, and compares the expected value of benefits with the expected value of costs. Flood control evaluation rests on this procedure. As discussed in the section on appraising flood control investments, many have questioned whether maximizing expected value of net benefits is an appropriate criterion. [See, for example, Kunreuther (1978), Heiner (1983).] Assigning probabilities to water flows is reasonably straightforward and the subject of a large literature. However, estimating probabilities for economic and political factors (prices, population, productivities) is only in its infancy. This, and the large computational load for serious probabilistic analysis has limited its application in practice to the flood control field.

Decision theoretic approaches [Dorfman (1962), Mishan (1976)] hold promise, but have seen little application beyond the theoretical level as yet.

A refinement of the expected value approach adopts Bernoulli's insight that individuals do not necessarily value uncertain prospects by their monetary expectation. The expected utility approach assumes that expectation is the appropriate decision criterion, but that individual attitudes toward risk should be reflected in the analysis. Risk aversion by individuals is hypothesized, so that highly risky projects have a certainty equivalent which is less than the expected monetary value [Dorfman (1962)].

Arrow and Lind (1970), however, have contended that it is not correct to presume that risk and uncertainty in benefit and cost streams are always socially costly. Depending upon (1) the entire portfolio of national investments, (2) the correlation between the benefit and cost streams of projects and the overall returns to the assets in the economy, (3) the existence of contingency claims markets, and (4) the extent to which the risk and uncertainty is spread over impacted individuals, it may be appropriate to ignore individual project risk in public project evaluation. The implication is that a risk-free discount rate should be used. Fisher (1973) has pointed out, however, that the theorem applies only to private goods, but not in the important cases where the goods provided are non-exclusive and non-rival (public goods). Pearce and Nash (1981) note some additional limitations.

The risk of dam failure is an aspect of water project appraisal that has been conspicuously ignored in federal water planning procedures. Experience has proven that dams do fail, so that siting dams above areas of large population must be studied with extra care. However, due to philosophic controversies over the concepts to be employed and the reluctance of water management agencies to admit publicly the possibility of failure, official planning guides do not yet address the issue. [See Baecher et al. (1980) for a discussion.]

3.5.5. Multiobjective appraisals

Conventional cost–benefit analysis has been criticized on the ground that economic efficiency was not the only criterion by which water projects should be judged. In a previous section, we have outlined the rise (and partial decline) of multiple objective planning procedures by the federal establishment in the United States. This development was paralleled by a burst of interest in formal evaluative procedures in the technical literature. Marglin (1962) and Freeman (1969) provided theoretical formulations incorporating additional objectives (such as regional income) into water planning models. Cohon and Marks (1975) Goicoechea et al. (1982), and Chankong and Haimes (1983) survey and appraise the various quantitative techniques proposed for formally solving multiobjective problems. Major and Lenton (1979) incorporate multiple objectives into a river basin planning exercise, while Gum et al. (1982) treat a water quality problem.

Relatively few economists have chosen to explore this topic. Many appraise it as merely an attempt to justify unsound public investments under the guise of broader criteria. (Some efforts in multi-objective evaluation have suffered from inadequate care in specifying objectives, identifying trivial physical impacts as their objectives rather than employing measures representing legitimate public goals.) Others doubt the ability of government "decision-makers" to provide quantitative tradeoffs among social objectives. Those economists holding this latter view tend to focus efforts on predicting the allocative and distributive consequences of water policy proposals, and recommend that the political system resolve the conflicts (as is provided in the federal planning procedures of the past decade). Haveman (1965), Gardner (1966), Infanger and Butcher (1974), and Miller and Underwood (1983), represent instances of measurement of distributive impacts of water-related programs and projects.

3.5.6. Ex post evaluation

The application of economic analysis to public spending decisions has its longest history in the water resources area. Yet, in contrast to many other program areas, analysis has been almost exclusively *ex ante* in nature. As a result of this singleminded attention to *ex ante* analysis, water resource planners have not had access to *ex post* information required to (1) determine if project evaluation techniques are biased, and if so, in what direction and to what extent, (2) revise evaluation methodologies so as to improve their reliability, and (3) gain information on the production functions on which project evaluations rest.

Several obstacles to meaningful *ex post* evaluation exist in the water area [Haveman (1972)]. The primary barrier, in addition to data problems, concerns the "with–without" framework of benefit–cost analysis. An *ex post* evaluation of the before–after sort is of no use in efforts to improve evaluation procedures. If, for example, the flood losses actually prevented by a project were estimated and used as a basis for judging the benefits produced by the project the appraisal would implicitly indicate that the prevention of damage to property induced into the floodplain by the project constituted a benefit attributable to the project.

Second, substantial conceptual and empirical problems are involved in appraising the performance of investment projects whose output depends on a stochastic process. When an investment has afforded protection against the occurrence of a probabilistic event, such as a flood, it has, in effect, a value that is analogous to insurance.

Finally, for most water projects, the bulk of expected project benefits occur in the later years of the project's life. In such cases, the analyst would find it most difficult to judge the efficiency of the investment on the basis of its output stream until it has matured.

In the primary water resource agencies the correspondence of actual costs with *ex ante* estimates has been explored periodically. In the 1950s a pattern of major cost overruns was uncovered; later cost analysis has been found to be far more accurate [Hufschmidt and Gerin (1970)].

Very few *ex post* appraisals have been undertaken. Haveman (1972) developed a framework for *ex post* evaluation of flood control, navigation, and hydroelectric project purposes, and undertook preliminary *ex post* analyses in each area. The study concluded that, "there is a serious bias incorporated into agency *ex ante* evaluation procedures, resulting in persistent overstatement of expected benefits" (p. 111).

Regional growth impacts. Another analytic thrust can be interpreted as a type of *ex post* appraisal. This is represented by the several studies which use econometric techniques to test the hypothesis – fundamental to much of public water policy – that water resource development does in fact yield significant regional economic growth impacts. A number of such analyses have utilized cross-sectional census data from one or more census periods and regressed various indices of growth (income, production, etc.) against expenditures for water development projects. Lewis et al. (1973) review the conceptual issues. Both Cichetti et al. (1975) and Fullerton et al. (1975) studied multi-state regions in the southwestern United States, but failed to identify statistically significant regional growth measures for federal irrigation projects. Howe (1976) surveys both the econometric and descriptive evidence from river basin developments on several continents, reporting rather mixed findings. The evidence at hand suggest caution is advised in projecting large and assured regional growth impacts from water projects.

3.5.7. Optimal sizing, timing and sequencing of water supply projects

This section calls attention to the related issues of optimal timing and sizing of projects. The sizing of water projects has often been based on hydrologic and engineering considerations. Size has tended to be made as large as possible, until the point where incremental cost rises sharply. This, however, ignores the likelihood of diminishing marginal net benefits. The conditions for optimal sizing of a project in a static and deterministic framework are achieved by the familiar equating of marginal costs and marginal benefits [Marglin (1962), Herfindahl and Kneese (1974)].

In the face of growing demand for project output, net present value of returns may actually be increased by postponing a project even though it might be found feasible for an immediate start. This will be the case if, roughly speaking, net benefits are growing at a percentage rate larger than the rate of interest. Herfindahl and Kneese (1974, p. 202) discuss the issue and Gittinger (1982, pp. 381–383) describes some simple tests for determining when to start a project.

Marglin (1963) presented one of the earliest formal attempts to deal with the problems of investment sequencing. Most of the recent contributions have come from the engineering and operations research literature. Moore and Yeh (1980), for example, apply a dynamic programming approach to a case study in which price-sensitive demand, a marginal cost pricing rule and an objective of economic efficiency are incorporated into a model for expansion of regional water supply capacity.

4. Evaluations of selected public water policy issues

Local, regional, and national government entities are involved in numerous water supply and management programs. We cover only the most significant in terms of public expenditures in the United States – natural hazard mitigation, water transportation and irrigation.

4.1. Natural hazards: Floods and droughts

Extreme climatic or meteorological events, such as floods or droughts may cause significant economic damages and have been a major preoccupation of water-related public policies.

4.1.1. Structural approaches to floodplain management

Provision of flood control is one of the major water resource activities in the federal government's portfolio. These activities often involve levee construction, channel improvement, and reservoir construction. In these cases, the purpose is to erect some man-made structure which will confront natural probabilistic events – the chance of a discharge of x volume occurring in any given year in any flood plain – and alter the impact of these events when they occur.

Several outputs from flood protection projects can be distinguished:

1. the reduction of crop damage from flooding;
2. the reduction of property damage from flooding;
3. the reduction of non-crop output losses due to flooding; and
4. increasing in the productivity of land and improvements on the floodplain.

Optimal flood control programs will minimize the sum of the costs of the protection program plus the costs of damage avoided. In Figure 11.1 the steps required to empirically estimate the present value of both crop and property damages averted are depicted and related sequentially to each other [see also James and Lee (1971), Haveman (1972), Herfindahl and Kneese (1974)]. These components of the empirical estimation process correspond to the method utilized

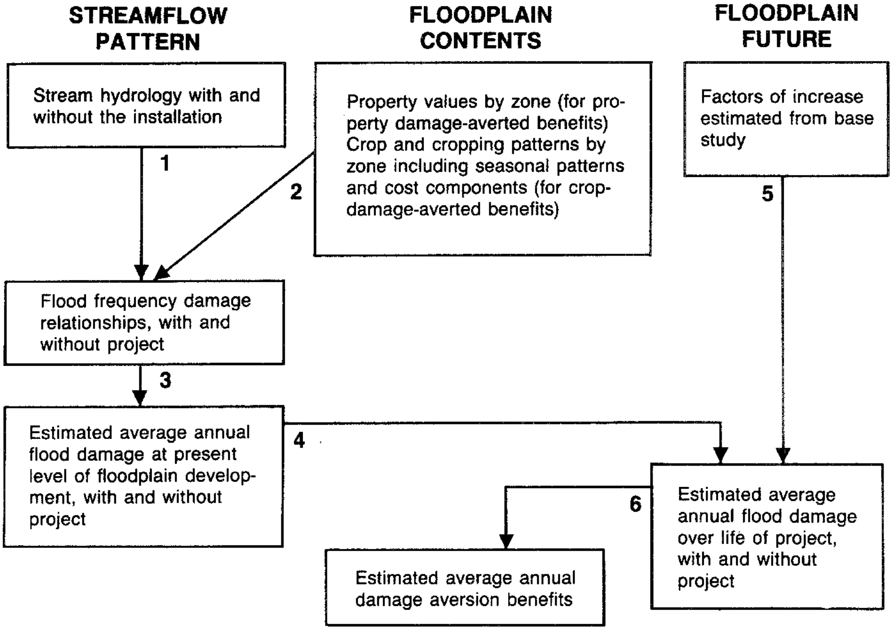

Figure 11.1. Framework for estimating flood control benefits.

by federal water resources agencies. As can be inferred from the figure, analyses of hydrologic conditions, forecasts of the physical performance of the installation, estimates of crop distribution patterns, seasonal planting patterns, flood-free crop yields and value components, and forecasts of "factors of increase" without the project are essential components of the estimated annual benefits resulting from the reduction of flood damage to crops due to the flood control project. Similarly, estimation of the benefits of averting damage to property required estimates of stream hydrology, property values of the floodplain, project performance, and factors of increase without the project.

Within a smoothly functioning economy operating at full employment, there are several additional benefits from flood protection investments beyond crop and property damage aversion. These include the reduction in output losses due to destruction of a crucial input or the temporary reduction or cessation of production. While these benefits are not properly considered direct project outputs, a net willingness to pay for them does exist. These benefits result from a reduction of production losses associated with activities not experiencing direct physical damage from flooding.

One of the outputs of a public flood control investment may be an increase in the productivity of floodplain land because of the reduced incidence of flooding.

For example, because of reduced flood incidence, an entrepreneur may find it profitable to shift floodplain land from low-net-yield pasture to high-net-yield agricultural commodities or to a factory site. In principle, this output is the net reduction in cost (or increase in net earnings) experienced by the occupants of the land because of the flood reduction services of the investment other than flood damage reduction. In the water resource literature, the value of this physical output is known as "land enhancement benefits".

Estimation of benefits in this productivity improvement category is difficult. In this case, the analyst must identify the increment to, and the composition of, economic activity that would have taken place on the floodplain in the absence of the investment, and compare this value with the actual increment to, and the composition of, economic activity occurring on the floodplain due to increased real productivity of floodplain land. Through this procedure, the investment-induced productivity increase of floodplain land can be estimated. Stated alternatively the analyst's task is to distinguish between (1) changes in the level and composition of economic activity on the floodplain attributable to productivity increases of floodplain land induced by the public investment, (2) changes in economic activity resulting from the natural growth of the area without the project, and (3) changes in floodplain economic activity attributable to inadequate or erroneous information about the extent of increased protection afforded the floodplain by the investment. Having isolated that component of changed economic activity attributable to investment-induced increases in land productivity, the analyst must estimate the increase in net earnings of these induced activities [see Lind (1967)].

These procedures for valuing flood protection benefits are fraught with both practical and conceptual problems. Here, only the primary of these will be mentioned.

(1) Flood protection benefits are "public goods" requiring demand functions to be added vertically.

(2) The willingness to pay benefits are based on the assumption that floodplain occupants are rational and fully informed. Substantial evidence exists to suggest that they do not act on the expected value concepts which dominate this area [see White (1964), Kates (1970)].

(3) If floodplain decisions are not made on this basis, non-optimal floodplain usage will result, both without but especially with flood protection. This will result in actual damages (and damages averted) being quite different than those predicted on the basis of expected values. (This perception has led to a de-emphasis on structural protection measures, and more emphasis placed on land use control, or mandatory flood insurance.)

(4) The benefits from land value enhancement have been disputed by several economists, who have argued that, generally speaking, land quality is a continuum, and that because land quality equivalent to that on the floodplain is available, land enhancement benefits are non-existent [Lind (1967)].

(5) It has been claimed that projected changes in land value can capture all of the benefit components described above. This is true, but only under a very restrictive set of conditions, including competitive markets in long-run equilibrium, no shifting of benefits outside of the floodplain, and fully informed floodplain decision makers operating on unbiased estimates of expected values [See Haveman (1972)].

Recent contributions include Cochrane's (1981) survey of the state of the art of flood loss simulation. He examines the potential for using these methods in developing nations. Cochrane and Huszar (1983) studied urban storm drainage practices, finding in a case study that application of engineering rules of thumb (e.g. "protect to the 100 year event") yielded excessively costly mitigation programs. Milliman (1983) surveys the economic literature on flood hazard mitigation and sets out priorities for economic research.

4.1.2. Non-structural measures for floodplain management

White (1964) led in advocating a shift away from strictly structural solutions to floodplain problems. Krutilla (1966) proposed flood insurance to shift the cost of risk-bearing toward the individual floodplain occupant. Floodplain zoning, flood forecasting and post-disaster relief policies have received extensive discussion, in light of the failure of the federal flood insurance program to live up to expectations [U.S. Congress (1979)]. The theoretical model of expected utility maximization hypothesizes risk-aversive behavior and suggests that affected individuals would be willing to pay more than an actuarily fair premium to protect against loss. Kunreuther (1978) demonstrates, with field and laboratory studies, that for relatively rare events which may cause high losses, such as floods, this hypothesized behavior is not observed. [See also Heiner (1983).] This finding may support compulsory insurance and regulations of land use and construction practices [Sorkin (1982)].

4.1.3. Drought impacts and their management

Drought is a condition of moisture deficit sufficient to have an adverse effect on plants, animals and ultimately man [Warrick (1975)]. Water managers attempting to mitigate drought impacts may modify demand, augment water supply, or select a combination of the two approaches. An optimal plan minimizes the sum of mitigation costs and drought losses. As with floods, the economic literature stresses the problems of measuring economic damages from moisture shortages, and treats both structural and non-structural approaches to drought mitigation.

Musgrave and Lesueur (1973) questioned the usefulness of large dams as drought mitigation measures in Australia, and proposed market-like procedures which would encourage scarce water supplies into the highest valued uses. The research agenda developed at the Harvard Water Program [Maass et al. (1962), Hufschmidt and Fiering (1966)] brought the field of stochastic hydrology together with economic analysis, and initiated concern with formally measuring loss or damage functions due to water shortages. Russell et al. (1970), developed two sets of loss functions: one corresponding to an *a priori* model of drought impact; and the other based on an *ex post* estimate of actual damages, both for the Massachusetts drought experience in the 1960s. Millan (1975) extended this tradition in a theoretically and empirically rigorous study of the upper Colorado River Basin in the western United States. An interindustry model of the economic region was adapted into a linear programming format to measure both direct and indirect economic impacts of variations in water supply. The major water-using sectors were each represented by several production activities designed to represent a range of adjustment to water shortage. The overall simulation model, in addition, incorporated advanced procedures for modeling the hydrologic system of the basin.

A number of studies cited in the subsequent sections on crop irrigation and on allocative institutions reflect attempts to model the impacts of water shortages [Anderson and Maass (1971), Dudley et al. (1972), Daubert et al. (1984), Angelides and Bardach (1978)].

4.2. Water transportation

Public provision of right-of-way and public improvements to existing rights-of-way for waterborne commerce are important water resource investment activities. They can in concept, be accurately evaluated by the "least-cost alternative" technique. The partial equilibrium presumption is that the nation requires a certain origin–destination movement of goods, and that the objective is to achieve this movement at the lowest resource cost to the nation. (The fixed origin–destination presumption, of course, is questionable.) The provision of transportation infrastructure, and the pricing (cost-recovery) procedure which accompanies it, will cause shifts in the location of both producers and consumers and, hence, changes in both the origin–destination pattern and the total volume of commerce requiring transportation services. It is generally accepted that analysis using such full general equilibrium approaches is not currently feasible.

This "lowest resource cost to the nation" concept can be elaborated still further in the case of, say, a waterway improvement, by introducing the distinction between the cost savings on an existing volume of traffic which is to be moved in a region *and* the cost savings on incremental traffic which is generated by the lower effective transportation *charges* in the region due to the introduction of the improved facility.

With this distinction, the total efficiency benefit of a navigation improvement is equal to the sum of the reduction in costs of moving existing traffic and the *net* willingness to pay for the additional traffic attributable to the improvement (the total willingness to pay for the incremental traffic less the incremental costs of accommodating that traffic, whether it is carried by water or an alternative mode) [see Haveman (1972)].

Current agency practice in the ex ante estimation of navigation benefits has been determined by direct congressional action in the Department of Transportation Act of 1966. This is the only project purpose for which Congress has explicitly dictated the definition of benefits and the concepts to be used by agencies in evaluation efforts. Largely because of this intervention, current navigation evaluation procedures deviate more from ideal procedures than in any other project purpose.

The early waterway benefit evaluation procedures utilized by the Corps of Engineers have been described by Eckstein (1958). The pre-1960 Corps practice was to evaluate the unit benefits of a navigation improvement by comparing the current rates that shippers would pay to transport commodities on the improved waterway with the rates they would pay for the next best alternative mode. Eckstein demonstrated that, because of the complex nature of the railroad rate-making process and the setting of railroad rates to cover full costs, it is the unit savings to shippers that are being estimated and not national resource cost savings per unit of traffic moved.

In estimating the volume of traffic that would utilize a proposed waterway, the Corps employed surveys of the commerce flowing into and out of the region. On the basis of the surveys, the reliability of which has often been challenged [see Nelson (1969)], an estimate was made of the volume of future traffic that would move by water. In practice, a sizable share of expected traffic growth in the region was often credited to the construction of the waterway traffic.

In this procedure, it was implicitly assumed that the difference between rail rates and barge rates would not change during the life of the project. Both potential technological change in the railroad or barge industries and competition from the waterway are ignored. Because post-waterway railroad rates are likely to decrease below their pre-project counterparts by more than barge rates for both of these reasons, there is sound *a priori* reason to expect that the traffic projected on the waterway by this technique is overstated. Prior to 1960, then, the procedures used to estimate both traffic on the waterway and unit savings on this traffic led to bloated estimates of the benefits from navigation improvements.

In 1960, a significant change in procedure was adopted by the Corps. The use of transportation rates was dropped, at least in concept, in favor of a comparison of the resource costs of shipping commodities by water with the costs of transporting them by an alternative mode in the absence of the project. However, even after the 1960 change, estimates of traffic expected to move on the waterway remained based on a comparison of current rail rates and barge rates expected to prevail when the waterway was completed. While the new cost-based procedures were superior in concept to pre-1960 procedures, in practice the reported estimates were similar under both methods.

In October 1964, however, a second revision in evaluation procedures was announced. This alteration – which took the form of an interim procedure to replace the post-1960 cost-basis-with-loopholes – was a substantial one in both form and practice. Proxies for long-run costs were used both in projecting traffic on the waterway and in measuring the benefits attributable to that traffic.

The effect of implementing this procedure was that estimates of traffic expected to move on the improved waterway were lower than estimates generated by earlier procedures. Because of the changes, fewer projects were able to demonstrate a benefit–cost ratio above unity. Congressmen and senators from states with strong waterway interests found this interim procedure to be a severe obstacle to project approval [see Haveman and Stephan (1968)]. Through Section 7 of the Transportation Act of 1966, Congress, led by the waterway interests, eliminated the interim procedure. The essence of this legislation was to force the Corps of Engineers to revert to the pre-1960 practice of estimating both waterway traffic and unit savings on the current rate basis.

As a result of this legislation, navigation benefits are based on an estimate of future waterway traffic, which, on *a priori* grounds, is seriously overstated, and on an estimate of unit benefits for this traffic, which represents savings to shippers rather than the appropriate (and smaller) savings in national resources devoted to transporting commodities.

As with other public water programs, cost-sharing has been an important issue in federal waterways policy in the past decade [Hanke and Davis (1974), U.S. National Water Commission (1973)]. The policy that the navigable waters of the United States should be free of tolls became explicit national policy a century ago. The conditions which may have justified a toll-free policy have changed, and in 1981 a tax on fuel used by shallow-draft inland waterway barge firms was initiated. Proponents of a tax cited familiar equity and efficiency arguments, while opponents were concerned that the competitive position of the barge industry would be weakened *vis-à-vis* the railroads, and that major waterways users such as agriculture, would be adversely affected. Shabman (1982) provides a discussion of these issues and the related literature.

4.3. Crop irrigation

Irrigation of agricultural crops represents the largest consumer of water diverted for human uses in the United States [Frederick and Hansen (1982)] and throughout the world. Particularly in desert climes, new water development can yield large increases in agricultural output, can stabilize agricultural returns, and promise to reduce unemployment and to stimulate economic development in rural regions. Economic analysis comes into play because both private and public development of irrigation water supplies can involve expenditures of large amounts of scarce resources, including capital, land and skilled labor. As the best sites become converted, the net economic benefit per unit of monetary outlay tends to fall, and many recent proposals do not stand up under careful economic scrutiny. Nevertheless, new irrigation projects remain high on the political agenda in arid zones of developed as well as developing nations. [Carruthers and Clark (1981) broadly treat irrigation economics.]

4.3.1. Productivity and demand for irrigation water

Numerous complexities are encountered in measuring the demand for irrigation water. First, crop production, with or without irrigation, is a biological process carried out in uncontrolled and highly variable environments. The process, therefore, is subject to the vagaries of diseases and pests and variations in climate (temperature, sunlight, wind, humidity, and rainfall), soil texture and fertility. Second, yield response to irrigation water application is especially sensitive to the rate at which water is combined with other inputs, such as soil nutrient levels and investments in on-farm application systems. Third, crop response may be inhibited by dissolved salts (salinity) in the irrigation water which become concentrated in the crop root zone by the evapotranspiration process. Finally, a realistic model wil reflect the fact that the productivity of irrigation water varies widely over the year, depending particularly upon soil moisture level and upon stage of growth of the plant.

The conventional approach to measuring the productivity value of irrigation water employs residual imputation on a crop-by-crop basis [surveyed by Young et al. (1972)]. More sophisticated approaches recognize the multi-crop nature of the typical farm and the sequential decision involved in scheduling the amount and timing of irrigations throughout the growing season. Moore (1961) in the United States, Flinn and Musgrave (1968) in Australia, and Yaron and his associates [Yaron and Dinar (1982)] in Israel were among the pioneers in formulating rigorous approaches to modeling aspects of the irrigation water allocation problem. See Vaux and Pruitt (1983) for an extensive survey of this literature.

The marginal value of irrigation water varies widely, due to the site-specific characteristics of production described above (climate, soils, crops, technology). They also may differ according to the conceptual framework employed in making the estimate (short run versus long run; public versus private accounting stance). Most surveys of the value of irrigation water have concluded that the long run average value is lower in irrigation than in competing offstream uses [e.g. Howe and Easter (1971), Young et al. (1972)].

Demand and price elasticity estimates cannot be derived by normal econometric procedures, due to lack of market exchanges. Howitt et al. (1980) summarize the parametric mathematical programming approach and argue for non-linear objective functions. National inter-regional programming models have been developed to analyze a number of issues related to irrigation water demand and to pollution from agricultural sources [Christensen et al. (1981)].

4.3.2. *Feasibility appraisals of irrigation investments*

Numerous critics of federal appraisal practices have perceived a general pro-development bias in official evaluation procedures. These critiques begin with Teele (1927), and include Eckstein (1958), Freeman (1966) and Young (1978). In addition to the general sources of bias discussed earlier (including using too low a discount rate, recognizing secondary benefits, ignoring the opportunity costs of water and failing to consider potentially less-expensive, perhaps non-structural alternatives) several issues have been raised which are specific to the irrigation case. The federal procedures for shadow pricing inputs and products have been challenged. A major objection is that no charge for the opportunity cost of family labor and management has been included. Also, project revenue forecasts tend to overplay the scarcity of future food supplies. Forecasted crop shadow prices have generally ignored the historic tendency of falling real agricultural commodity prices even though persistent excess capacity in the U.S. agricultural production plant has led to expensive federal government supply control programs. Social benefit estimates should be net of subsidies to maintain crop prices [Martin (1979)]. Another bias has been the overweighting of the projected crop plan in the direction of high-margin specialty crops. At least a portion of the high margin in such crops should be charged against specialized management and risk-bearing, rather than credited to irrigation water. Since new production from irrigation development will, at the margin and from the national perspective, largely be registered in forage and feed grain crops, the high income specialty crops should have minimal weight in project appraisal. (Many of these criticisms have been addressed by the U.S. Water Resources Council's 1983 planning procedures but the Congressional practice of "grandfathering" projects appraised under earlier, less rigorous standards delays the effectiveness of such reforms.) Dudley et al. (1972) analyzed several important planning issues, including the optimal proportion of water to land. Martin (1979) and Sampath (1983) extended the typical partial equilibrium approach to benefit measurement, and appraised irrigation programs in a welfare framework which employs consumer as well as producer surpluses in measuring irrigation benefits.

External costs, such as the reduced productivity in downstream regions from saline irrigation return flows must be recognized [Moore (1981), Oyarzabal and Young (1978)]. In many cases, irrigation project plans must anticipate costs of drainage to alleviate downslope waterlogging and salinization from irrigation developments [Moore (1972), Johnson (1981), Carruthers and Clark (1981)].

Finally, insufficient attention may be given to assuring that the proposed public investment is the least-expensive means of achieving the same project outputs.

While lesser-cost alternative may be structural, such as reliance on private groundwater development in lieu of public surface-water projects, it might be non-structural, as when a quasi-market system is introduced to ration water supplies.

The equity and efficiency issues involved in federally-subsidized irrigation water supplies aroused debate in the 1970s [U.S. National Water Commission (1973)]. The "160 acre-limitation" in the 1902 Reclamation Act supposedly limited the subsidy to family farmers, but changing technologies and the possibility of significant economies of large size created a thrust to relax the limitation [Martin (1978)]. Seckler and Young (1978) challenged the notion that significant economies of size existed beyond 160 acres [see also Moore (1982)] and criticized the distributive impacts of the then-existing administrative procedures. They suggested that water, rather than land be the basis of the limitation and proposed a two-tier charge system in which subsidized water would be provided up to a specified limit and full cost charged to farms beyond that point. Legislation approved in 1982 raised the ownership limit to 960 acres, while providing for full cost water charges for farms larger than that cutoff.

5. Further topics relating to water supply

The previous sections dealt with theory and practice in evaluating large scale surface water impoundments by public sector agencies and related issues. Next, several aspects of water supply are taken up. Space constraints prevent coverage of such special topics as desalination and waste-water recycling.

5.1. Groundwater management

An aquifer is a geologic formation of permeable materials which is saturated with water. The earth's groundwater resources are extensive – they constitute probably the largest source of fresh waters. Two hundred billion acre feet underlie the land surface in the coterminous United States [U.S. Water Resources Council (1980)].

The groundwater resource presents problems of quantity and quality similar to surface water as well as a number of special issues. The major characteristics of aquifers that affect the costs of water supply are (a) depth to water table, (b) thickness of saturated zone, (c) transmissivity (the rate at which water is transmitted through the aquifer), and (d) characteristics of the formations through which the well must be drilled [Carruthers and Clark (1981)]. Much of the economic literature can be interpreted as incorporating notions of diminishing returns, time discount and external costs into a field still dominated by simplistic "safe yield" concepts.

Groundwater basins are typically exploited by a large number of independent pumpers withdrawing from the common groundwater supply. The economies of large size so evident in surface water developments are not present in groundwater extraction. Groundwater systems vary greatly in their physical characteristics, and obtaining information on the depth, porosity, and water quality in an aquifer can be a costly undertaking. As a result, accurate representation of the physical system in order to predict impacts of alternative extraction policies may be difficult to achieve. [See Gorelick (1983) for a survey of physical aquifer models.] As groundwater ordinarily can move (slowly) in response to withdrawals, the action of any one pumper affects the conditions experienced by other users; thus they are interdependent and external costs (or benefits) are imposed.

Under unregulated management, where non-renewing groundwater is held in common ownership and utilized by otherwise independent agents, the resource is "fugitive" and must be captured in order that the user can claim property rights to it. The individual user's property rights to future use of the pool are indefinite, as other pumpers may utilize the water in the meantime. In such instances, the self-interest of the individual user may lead to socially non-optimal pumping regimes. This "common pool" problem is conceptually similar to that of the open access oil pool or fishery, which has been extensively analyzed [Haveman (1973), Ciriacy-Wantrup and Bishop (1976), Randall (1983)].

O.R. Burt, in a series of influential papers, contributed fundamentally to the understanding of optimal groundwater management. He incorporates both direct and external costs of pumping and the tradeoffs between diminishing returns to usage now and the present value of future uses into a dynamic optimization framework [see Burt (1964, 1975)]. Brown and McGuire (1967) also employed a single-celled aquifer model, developing an optimal policy for allocating publicly-supplied surface water conjunctively with groundwater. Bredehoeft and Young (1970) treated a multicell aquifer case, and as did Mapp and Eidman, (1975) developed more realistic representations of crop response to reduced irrigation water supplies. Daubert et al. (1984) treat a somewhat different issue, that in which heavy groundwater withdrawal from a renewable tributary aquifer reduces stream flow, and adversely affect supplies available to senior surface water rights holders.

Finally, withdrawals of groundwater often impose spillover costs, i.e. encouraging flow of poor-quality water into parts of the aquifer [Cummings (1971)] or subsidence of overlying land surface [Warren et al. (1975)].

5.2. *Interbasin water transfers*

When the origin of transferred water is in a different hydrologic region or basin (and often a different political jurisdiction) than the destination of that water, a major conveyance system is typically required. The change of ownership rights can involve interstate (or province) or even international institutional considerations and conflicts. Inter-basin Water Transfer (IBWT) proposals typically require funding and political sanction from the national government. They raise serious economic problems, particularly those involving the perceived demands of the importing basin versus the possible future needs of the basin of origin. Income distribution and environmental impacts and the effects on communities and associated sociopolitical institutions are other concerns.

The most general treatment of the economic considerations for evaluating an IBWT is provided by Howe and Easter (1971), who advocate the need for rigorous economic analysis of primary and secondary benefits and costs from both a regional and national viewpoint. Adapting E.N. Castle's suggestions, Howe and Easter set out two principal conditions for economically efficient transfer of water which can be expressed as follows:

(a) The increments to net incomes in the importing region or regions must exceed the sum of (i) the loss of net incomes in the exporting region, (ii) net income losses in regions whose outputs are competitive with those in the importing region, and (iii) the costs of the physical conveyance systems.

(b) The cost of the physical transfer system must be less than the cost of the best alternative for supplying the same amount of water to the importing region.

Net incomes and costs are assumed to be correctly expressed in present value terms on the basis of a consistent time period and discount rate. The calculation of net incomes should include direct impacts, real external costs (water quality degradation, forgone power or recreational benefits) and real secondary benefits which may arise from a departure from the competitive conditions including unemployed resources, immobility of resources and the existence of economies of large scale in production. Bain et al. (1966), Hirshleifer et al. (1960), Hartman and Seastone (1970), Kelso (1973), Cummings (1974), and Supalla et al. (1982) have studied particular cases.

5.3. Conservation

Water "conservation" has been suggested as a policy tool for managing increasing water scarcity, so much so that the concept became a key element in the Carter Administration water program. A difficulty is that, as Mann (1982) notes, several different meanings can be attached to the term. Therefore, while "everyone is in favor of 'conservation', no matter what it means" the idea has some practical limitations.

Generally, conservation means an avoidance of wasteful usage, but one person's waste often is another's benefit. Those who perceive the value primarily in developing and withdrawing water for human consumption and production activities tend to regard water which is not "used" in this sense as wasted. This view, in the extreme, holds that water is wasted if it flows by a potential dam site or to the sea.

A much different position emphasizes the aesthetic value of water in its natural, free-flowing state. Some of the more utilitarian off-stream uses of water are wasteful from this perspective, while conservation is understood as protection from such "less valuable" uses.

Yet another concept of conservation is improving the technical efficiency of water use. Losses, due to leakage, evaporation or avoidable wastage in production and utilization of water should be reduced where technically feasible [see, for example, Flack (1981)]. This view, however, may not carefully consider the economic costs of reducing wastes relative to the gains. The "conserving" of one resource will usually imply use or depletion of one or more other resources. Also, in many instances, losses to an upstream user are the downstream user's supplies, and technically efficient solutions may have unexpected basin-wide implications.

The economic approach to conservation [Ciriacy-Wantrup (1952)] views conservation as an economically efficient allocation of resources encompassing the dimensions of time and space. This perspective emphasizes that the opportunity costs of other resources, in addition to that of water, must be considered in assessing waste. Pricing water at its marginal supply cost (incorporating the opportunity cost of water itself) will assure that water is not wasted, since the rational user (private or public) will not over- or under-invest in water supply capacity when faced by the appropriate incentives [Griffin and Stoll (1983)]. Where markets and pricing are not feasible, benefit cost analysis, employing shadow prices for non-marketed impacts, will aid in preventing wasteful use of water as well as of related resources. Saving water is truly "conservation" in the economic sense only if the benefits of the water-saving technology exceed the costs.

5.4. Water supply from reallocation

A regional water economy can be characterized as being either in an "expansionary" or a "mature" phase [Randall (1981)]. In the expansionary phase, the incremental cost of new water supplies remains relatively constant (in real terms) over time, and water development project sites are available to meet growing demands. The mature phase is characterized by rapidly rising incremental costs of water and increased interdependencies among water uses and users.

The rising cost of water supply in a maturing water economy brings about a search for sources of water among existing uses whose incremental value productivity is less than either the cost of new supplies and the benefits of new uses. Since crop irrigation typically accounts for 80–90 percent of water consumption in arid regions, reallocation of water from agriculture to sectors with rapidly growing water demands is receiving increasing attention.

Most surveys of the value productivity of irrigation water [e.g. Howe and Easter (1971), Young et al. (1972)] have concluded that the long run marginal value is lower in irrigation than in competing offstream uses or than in some instream uses. Put another way, the maximum willingness to pay in agriculture tends to be much less than the willingness to pay for water by households and industries. Hence, where municipal and industrial demands are rapidly growing in water-scarce regions, forgone net benefits from reducing agricultural use may be less than the costs of a new supply. Substantial economic savings can be achieved from reallocation to the higher uses as compared to constructing new water supplies [Kelso, Martin and Mack (1973)].

The above hypothesis has proven controversial in some quarters. Arid-region governments have exhibited special concern for both the farm water users and the forward- and backward-linked economic sectors supplying inputs, processing and marketing services. The conventional wisdom has held that the indirect effects of reallocation on employment and income would be large, such that the full costs of removing water from crop production would be unacceptable.

The empirical evidence seems to suggest otherwise; that the economic impacts would be relatively limited [Young (1984)]. Water removed from irrigation would be the least valuable, drawn largely from the food and feed grain and forage sectors. Since foreseeable urban growth would account for only a small percentage reduction in irrigation water supplies, the sacrifices in net productivity would be minor relative to the gains in the growing sectors. These sectors also account for relatively small indirect incomes per unit of water consumed as compared with those from the emerging urban sectors. Also, inexpensive water may be obtained by reducing seepage in irrigation canals [Stavins (1983)].

State water and property laws generally protect the interests of farmers whose water is demanded by urban sectors; indeed, they often reap large capital gains in the transfer. The rate of loss of irrigation water, even in highly urbanized areas, will be slow, on the order of one to two percent per year. In such cases, the indirectly affected workers and businesses have time to anticipate and adjust.

6. Institutional arrangements for allocating and pricing water

This section treats the issues of allocating water among users and of pricing as largely separate problems. Note, however, that the literature outlining an ideal water market system and that dealing with water pricing appear to converge with the notion of a pricing system which reflects the opportunity costs of water with the mechanism of transferable water rights or entitlements.

6.1. Water allocation institutions

The institutions that affect the allocation of water among competing uses are crucial in determining the efficiency and equity effects of water use. These institutions concern "...sets of ordered relationships among people which define their rights, exposure to the rights of others, privileges, and responsibilities" [Schmid (1972, p. 893)]. These rights – basically, property rights – structure the incentives and disincentives between and among individuals in their decisions regarding water use [Ciriacy-Wantrup (1969)]. In the United States, these institutions are largely established by the individual states.

The research and writing concerning the structure of water allocation institutions is a part of what has become known as "analytical institutional economics". These writings are both positive and normative, and consist of the application of the neoclassical micro-economic research program to the laws, rules, and other institutions affecting water allocation. Ciriacy-Wantrup contributed prominently and his (1967) essay summarizes much of his thinking. The exhaustive studies of water rights law [e.g. Clark (1967)] are indicative of the complexities involved.

6.1.1. Goals of water rights systems

The selection of a system of water rights involves a compromise among several, often conflicting, social goals. Not surprisingly, therefore, different cultures have chosen different forms of water institutions, reflecting the relative importance of the various objectives. In the United States, the principal objectives which a water rights system is expected to attain are often stated as economic efficiency and fairness (i.e. equal treatment of equals). In other cultures, however, the desire for orderly conflict resolution and for popular participation and local control are of at least as much importance in understanding the evolution of water rights systems [Maass and Anderson (1978)].

An important factor influencing the form of water institutions in a society is the relative scarcity of water. A second factor is the transactions costs required to establish and enforce a water rights system. Where water is plentiful relative to demand, laws governing water use and allocation tend to be simple and enforced only casually. When water is scarce, however, more elaborate systems of rights have evolved. In many regions, water supplies are only now becoming sufficiently scarce to require more formal allocation mechanisms. The resulting conflicts could be mitigated if a well-defined market system for transferring rights was available in which compensation could be easily provided for the reallocation of rights. Institutional innovations to create such mechanisms tend to emerge in response to increasing water scarcity and reduced cost of enforcement [Ruttan (1978)].

6.1.2. Desirable attributes of a water allocation system

The attributes of a water rights system are closely linked to the objectives which the system is expected to attain. Nearly all of the literature concerned with designing water allocation systems have focused on those attributes required for achieving allocative efficiency. Two attributes associated with this objective are "security" and "flexibility" [Ciriacy-Wantrup (1967), Trelease (1965)]. A system is secure if it affords protection against legal, physical and tenure uncertainties. Only when rights are reasonably secure, will users undertake profitable long-term investments to capture and use water. Flexibility refers to the ability to change at low cost the allocation of water between regions, uses, and users over time – in short, the ability to accommodate changes in demand, reallocating water to higher-valued uses as they emerge. While security is desirable, it can cause a reduction of flexibility.

A third attribute is certainty – the rules of water use must be easy to discover and to understand. However, even if the rules are certain, basing them on the concept of water as a "free good" results in a rule of capture and the associated overuse and misuse of the resource. A final consideration, emphasized by Trelease, is that a desirable system should minimize the possibility that water users would impose uncompensated costs on third parties.

To sum up, a system of rights must be well-defined, enforced, transferable, and confront users with the full costs of their actions. An institutional arrangement with these attributes will permit the establishment of a market for rights which will readily reflect changing demands.

6.1.3. Empirical studies of water allocation institutions

Numerous studies have analyzed the nature of various water allocation systems, and the relationship between the characteristics of systems and the efficiency and equity impacts of water use. The issue of water transfers is a primary focus of these analyses, in particular, (1) the pecuniary and technological externalities associated with private water transactions and the importance of these in generating political opposition to large scale transfers [Hartman and Seastone (1970)]; (2) the forbidding of transfers in the presence of any such external effects (which implicitly presume that these spillovers are infinitely costly) which are present in water laws of most western states [Ditwiler (1975)]; (3) the consumptive use basis for the measurement of the benefits and costs of water transfers [Johnson et al. (1981)]; (4) the integration of the pricing and water allocation literatures and the potential of "transferable water entitlements" to facilitate water reallocations in water scarce economies with high incremental costs of new supplies [Randall (1981), Howe et al. (1984)]; and (5) the evaluation of potential gains from replacing the public sector constraints with allocation institutions having the characteristics of a market [Gardner and Fullerton (1968), Angelides and Bardach (1978), Brown et al. (1982), Howitt et al. (1982), T.L. Anderson (1983), Wong and Eheart (1983)].

6.1.4. Institutions for allocating irrigation water

Since irrigation is the largest consumer of water worldwide, and receives a major share of public investment funds devoted to water supply and management, institutions for its allocation deserve special mention.

The complexity of the institutional arrangements necessary to efficiently and equitably operate irrigation delivery systems is not generally recognized, even among social scientists. This complexity arises mainly from four sources: (1) the large number of individuals whose disparate water needs must be met; (2) the tendency to large seepage losses of water along pervious watercourses; (3) the physical interdependence among water users, such that one's water supply depends on the actions of others up-canal; and (4) the variability of water supplies due to natural climatic swings. These conditions require "institutional arrangements" in the sense used by Fox (1976), denoting an interrelated set of entities (organizations) and rules which serve to organize activities to achieve social goals. Organizations to represent the interests of the numerous water users, to establish and enforce rules and to maintain the ditches are an important institutional need in public irrigation systems [Easter and Welsch (1983)]. Many countries tend to provide the physical delivery system and assume that the local water management institutions will automatically follow.

Numerous forms of rules which allocate water among users are observed, ranging from informal systems where all farmers can take water whenever it is available to the elaborate volumetric pricing system found in Israel [Yaron (1979)]. The main limitations of existing systems are in apportioning seepage losses equitably from head to tail along the ditch and protecting downstream users from unauthorized withdrawals above them [Bromley (1982)]. Wade (1982) describes the tendency for lack of enforcement of established distribution rules, such that the wealthy and the powerful are favored. Seckler and Nobe (1983) emphasize the role of management controls in alleviating such problems. Maass and Anderson (1978) combine careful institutional descriptions of case studies in Spain and the United States with a series of simulation model analyses of the economic impacts of alternative sets of allocation rules. Bromley et al. (1980) recognized the parallels between the issues involved in the initial organization of a group of irrigation water users and the problem of the "Just Society" studied by the philosopher John Rawls. They set out a sophisticated set of rules and procedures for establishing water users' organizations.

6.2. *Beneficiary charges for water use*

This section focuses on the problems of beneficiary charges when water with private good characteristics is publicly supplied. Thus, the cases where water is supplied privately or through a market are not dealt with, nor are the problems of financing programs with non-rival, collective good attributes, such as those for water quality improvement or flood control. In the cases analyzed, the rates or prices set have both resource allocation and equity impacts, and influence the level of agency revenues. The principles of pricing or rate setting are taken up first, followed by the related issues of cost-sharing and cost allocation.

Empirical evidence on the effect of pricing on water consumption suggests that imposition of a measuring/metering system together with volumetric charges results in significant impacts on consumption. Schramm and Gonzales (1976) present a case study on irrigation in Mexico, while Hanke (1970) and Gysi (1980) report on the effects of water meters on urban consumption. Once a metered system is installed the demand has been found to be relatively price inelastic. [See Howe (1982) and C.E. Young et al. (1983) for cross-sectional estimates of price elasticity of household water demand.]

6.2.1. *Rate-setting*

Rate-setting represents a choice of policies within a multiple objective framework, in which (a) allocative (Pareto) efficiency, (b) equity of income distribution, and (c) "fairness" of apportioning costs are all major social objectives. Subsidiary criteria include simplicity, administrative feasibility, and stability [Bonbright (1961, pp. 290–292)]. A general principle or rule for setting rates can be associated with each major criterion. The principles each convert one of the major social goals into a broad practical guide or formula for setting rates.

The marginal cost pricing principle is the rate-setting rule applied where allocative efficiency (maximizing net social product) is the primary objective. When rates are set according to the schedule of marginal cost of supplying water, then the user will demand the commodity so long as marginal willingness to pay exceeds incremental cost, and the optimal level of usage will result. A corollary of this principle is that the common practice of "flat rate" pricing of water, in which no marginal charge is imposed, is likely to encourage consumption beyond the optimal level.

While economists have generally endorsed the marginal cost principle, application of it is difficult because of the variety of definitions of the appropriate marginal cost concept for pricing policy [Saunders, Warford and Mann (1977)]. An example concerns the transactions costs associated with measuring, allocating and monitoring a water pricing system. In an irrigation system with plentiful water supplies and numerous small field units, the transactions costs of a volumetric pricing system may exceed the value of water saved [Bowen and Young (1983)]. A second example is the long debate over the "Short Run Marginal Cost Principle" stemming from the work of Lerner and Hotelling in the 1930s. Strong objections to setting public utility prices equal to marginal costs, especially where marginal cost is below average cost (hence, requiring public subsidy) emphasized the absence of a market test to determine whether users were willing to pay the total cost of supplying the commodity, the potential misallocation of resources stemming from the additional taxation, the redistribution of income in favor of users of products of decreasing cost industries, and the impetus toward centralization of the economy [Coase (1971)].

While most of these objections can be dealt with by a multipart pricing system (involving price set equal to marginal cost, plus an assessment levied on users to reflect the costs which do not vary with output), establishment of such multipart systems which accurately reflect costs is difficult. Multipart rate structures are now frequently found in municipal and industrial, irrigation and hydroelectric power systems. Water pricing in Israel, in which all users are metered and face a rate structure with increasing blocks is an example [Yaron (1979)]. However, as they have been applied, multipart pricing systems often fail to account for an economically correct concept of opportunity costs, focusing rather on historical or embedded costs. The opportunity costs which are relevant include both the value of water in alternative uses or at a future time, which is called "user cost" [Howe (1979)]. Also relevant are the costs of securing incremental supplies in the presence of demand growth [Munasinghe and Warford (1982), Milliman (1971), Davis and Hanke (1971), Seagraves and Easter (1983)]. In this view, historical costs are sunk, and therefore irrelevant to establishing an efficient rate structure. Moreover, the opportunity costs of water should be determined by a market mechanism rather than by administrative procedures [Randall (1981), Howe et al. (1984)].

We turn next to a brief discussion of some alternative rate-setting principles which have been proposed or utilized.

The "ability to pay" principle is an alternative principle for rate setting, and rests heavily on the equity criterion. The rule provides the most common basis for setting rates for irrigation in the United States (and elsewhere), and is also regularly applied to village water supplies in developing countries. A common practice is to require only operating costs to be recovered fully plus a small fraction of the initial investment.

The U.S. experience with federal irrigation projects is illustrative. Originally planned early in this century according to a full cost recovery concept, three decades of unsuccessful attempts to fully recover costs ensued. In implicit recognition that costs overshadowed benefits (thus yielding zero demand at full costs) an ability to pay procedure was authorized in 1939 [Huffman (1953)]. A complex formula has been developed which limits the farmer repayment requirement to about 10–20 percent of estimated federal costs [North and Neely (1977)].

The ability to pay approach has little to commend it except in instances where low income groups are to be explicitly subsidized. The concept is relatively subjective, and political pressures arise to set the formula in ways which redistribute income from taxpayers to water users. Since charges bear little relation to costs, whether consumers would be willing to pay the total costs of supply is not tested.

The "net benefit" principle, sometimes termed the "rent" principle, seeks through charges to capture most or all of the economic surplus accruing to the user. Net productivity of the user would govern the calculation, but neither past or opportunity cost would enter in. The approach has been proposed for pricing public irrigation water, and is often embraced in more centralized political systems [Ansari (1968)]. The benefit principle is consistent with the view that water and its fruits are the property of the state. Setting rates strictly on the basis of net benefits appears to reflect a relatively deterministic view of the resource allocation process, one which ignores the incentive effects of pricing structures and appears to violate accepted equity principles.

The average cost principle calls for recovery of all costs by charging for each unit received according to the average cost. It is simple and easy to understand. It is fair and equitable in that beneficiaries pay just the resource costs incurred in their behalf. The desired signals to resource users are provided, although not in so precise a way as could be achieved by multipart pricing. As the approach is usually applied, however, historical or "embedded" costs serve as the basis of the calculation rather than opportunity costs.

In sum, in some places water is not yet sufficiently scarce to justify the tangible and intangible costs of establishing formal pricing systems. In such cases, flat rates will satisfy repayment requirements. However, when signals of scarcity of water (and of the costs of related construction capital and labor) are absent, pressures arise for structural solutions to satisfy incorrectly perceived water "needs". The expectation of increasingly scarce water supplies suggests adoption of entitlement and rate systems which reflect supply costs and the changing opportunity costs of water [Randall (1981)]. Such systems can be both efficient and fair, and have been observed to be operable in practice [Howe et al. (1984), Brown et al. (1982)].

6.2.2. Cost-sharing and repayment

The terms "cost-sharing" and "repayment" refer to the rules and procedures for apportioning the costs of federal projects in the United States among the federal government and local beneficiary groups such as irrigation or flood control districts, state fish and game or recreation agencies, or water quality control organizations. Since cost-sharing rules determine how and by whom costs will be borne, they represent the government's *de facto* pricing policy when the outputs of projects are not explicitly priced. The trade-offs among allocative efficiency, equity and fairness have come to a focus in the cost sharing rules actually adopted by the government.

North and Neely (1977) have shown that federal cost-sharing rules vary widely according to the agency constructing the project, the technology adopted to serve the various project purposes, the type of project output, and whether the costs in question are initial capital costs or operating and maintenance costs. Critics of federal policies have identified biases in existing rules such that the benefit-cost ratio calculated from a local perspective is substantially larger than that calculated from a national point of view, leading to inappropriate incentives [Allee (1982)].

Consistency in sharing costs between agencies, between structural versus non-structural solutions, between capital and operating costs has yet to be achieved. Davis (1968) examined the effect of repayment rules on the choice of water quality control techniques. Marshall (1970) and Marshall and Broussalian (1972) have extensively discussed conceptual issues. Marshall proposes adoption of the "Association Rule", which would require the local share of project cost be in proportion to the local share of benefits. Other rules which would tend to provide appropriate incentives include: uniform costs shares for all techniques (including non-structural); uniform cost shares for all cost categories; and uniform cost shares for all agencies for a given project purpose.

The issue only arises, of course, in the absence of a systematic long run marginal cost pricing system. However, it must be recognized that the traditions of pricing some water project services (i.e. irrigation water) on an ability to pay basis, and the difficulties in identifying beneficiaries of non-exclusive services (water quality improvement, flood abatement), create complex practical problems. Both the Carter and Reagan administrations have promised major revisions in cost-sharing policies, moving toward larger local shares and "up-front" contributions by states or beneficiary groups. As of this writing, no formal agreement among the executive branch, the Congress and the states on this issue has been announced.

6.2.3. Cost allocation

"Cost allocation" is the process of assigning an appropriate share of joint multiple purpose project costs to each project purpose or user class, and is a basic measurement issue in designing appropriate pricing or cost-sharing policies. User classes may be grouped according to economic sector, political sub-division or both, and joint cost allocations among them have both allocative and distributive implications.

Given the nature of the problem there is no ideal allocation procedure, and some degree of arbitrariness afflicts all of the suggested alternatives. Gittinger (1982, p. 233) and James and Lee (1971, p. 259) each list several guidelines for selecting allocation rules, of which three stand out. First, the method should be fair in that the user class be charged at least the incremental cost of receiving project benefits. Second, the joint cost allocation procedure should not make infeasible any service class for which incremental benefits exceed separable costs. Third, no class of service should be assessed charges in excess of the benefits to be received.

Numerous cost allocation formulas can be identified, the most common of which are the "Proportionate Use of Capacity" and "Separable Costs–Remaining Benefits" (SCRB) methods [James and Lee (1971, p. 533)]. Because the first method assigns joint costs in proportion to the quantity utilized, expressed in terms of volumes or flow rates, it may be difficult to apply in cases where project outputs cannot be measured in volume terms, as with non-consumptive uses, water quality, or flood control. A more significant objection to this procedure is that it can fail the second or third guidelines above [Herfindahl and Kneese (1974, pp. 291–292)].

The SCRB method allocates to each user class the identifiable (or separable) costs of including that purpose or service in the project, plus a share of the joint or common costs. The joint cost share is allocated as a proportion of the benefits net of separable costs ("remaining benefits"). The SCRB method satisfies the guidelines listed above, and is relatively simple to apply. Accordingly, it has been selected by federal agencies in the United States as the most acceptable approach. Loughlin (1977) proposes a refinement to deal with a possible inequity in the sharing of the savings resulting from multipurpose developments as compared with single purpose projects.

Some recent cost allocation proposals are based on a game theoretic framework. The theory of cooperative games provides approaches to joint cost allocation which take strategic possibilities into account. Heaney and Dickinson (1982) provide an integration of this literature with the more traditional analyses. See also H.P. Young et al. (1982) and Loehman et al. (1979) for applications. These highly formal approaches identify limitations of the traditional (i.e. SCRB) methods, but their complexity has inhibited the adoption of alternative solutions at the applied policy level.

7. Federal water management strategies

In large part because of the economic characteristics of water, a wide variety of institutions have grown up around the productive and consumptive uses of this resource. A simple listing of these arrangements will indicate their diversity and pervasiveness:

- to improve water quality a complex subsidy-rule enforcement strategy has been adopted by federal and state governments; construction of municipal waste treatment plants is highly subsidized by the federal government;
- massive public works expenditures have been undertaken to reduce flood damage, reduce transportation costs, reduce municipal and industrial water supply costs, provide irrigation water, and to expand recreation opportunities;
- publicly-imposed charged have been established for irrigation water, hydroelectric power, use of inland waterways, and recreation;
- administrative regulations for evaluating proposed public water resources investments have been established;
- regulations for using water resources for navigation, irrigation, and recreation purposes have been imposed;
- mandatory flood insurance legislation interacts with flood control investments, as do local floodplain zoning restrictions;
- provision for privately initiated litigation to prohibit public agency construction for rule enforcement, or to modify construction or regulatory plans, exists and has been widely used.

All of these institutional arrangements have evolved to deal with problems in the use of water resources, and most have been designed to deal with inefficiencies in water use which would occur because of the physical and economic character of water resources. Yet, in many cases the institutions which have been developed have serious problems of their own, and may generate their own inefficiencies. Indeed, in some instances it is not clear that the institutional arrangements in existence are superior to no public or collective intervention.

In the following paragraphs, we will indicate the nature of the inefficiency problems created by a few of these institutional arrangements, and in so doing indicate a variety of the reforms in water resource management which have been proposed [see also Lord (1979)].

7.1. *Water pollution control strategies*

The record of policy effectiveness in this area is mixed, at best. Subsidies to only conventional "end of pipe" treatment leads to a concentration of resources on this approach when a wide range of other, less costly, waste reduction techniques are available. The drive for technology-based standards and individual source permits fails to recognize differential watercourse capacities, different discharge levels in a region, and different levels of potential recreational use in a region, and tends to lead to excessive control costs. The subsidy funds have been dissipated by failing to concentrate them on municipalities with the most harmful waste loads and by restricting their use to only the construction of waste treatment facilities and not their operation. By subsidizing the capital costs of treatment facilities, sewer charges on industrial, commercial, and domestic dischargers are decreased, providing a windfall to polluters and decreasing their incentive to reduce discharge flows.

It is also generally agreed that the rule enforcement aspect of the strategy has been less than fully effective. Political bargaining is the very nature of the rule-making/enforcement approach and in this and other cases the bargaining is between parties of unequal power. While the regulatory process is often viewed as an instrument for public control over the behavior of the regulated, the opposite result has often occurred. At every stage in the bargaining process, those being regulated have much at stake while the public interest is diffuse, poorly organized, and represented. Predictably, the bargains struck favor the regulated. In the case of water pollution, the enforcement process has been long and drawn out and costly in terms of legal and administrative resources. Industrial polluters confront higher marginal returns from employing legal counsel to oppose and negotiate enforcement efforts than in undertaking pollution control efforts.

Policy measures without these adverse efficiency and equity consequences have been studied and proposed in the water pollution area. In particular, the case for the application of an effluent charge [Kneese and Schultze (1975)] or marketable effluent permits [Joeres and David (1983)] to water pollution control policy seems especially strong. Incentives for waste reduction would be provided, the windfall gains implicit in the current policy would be avoided, and the allocative inefficiencies generated by regulatory uniformity would be corrected. Perhaps most importantly, the system of bargaining between regulator and regulatee would be replaced by a system which rewards reductions in residuals discharges.

7.2. Public water resources investment policy

Several aspects of public investment policy in the water resource area have led to resource misallocation, inequity, and incentives for inefficient private sector behavior. The efficiency implications of a few of these have been discussed above; others have been well-documented in the literature.

The contention that project evaluation procedures have generally overstated some categories of benefits and understated costs was argued in Section 4 above. These biases are not likely to be fully corrected until the measurement and evaluation function is removed from the operating agencies and placed in the domain of an agency with independent evaluation capabilities. The agency chosen should not only be responsible for evaluating the efficiency and equity impact of proposed projects, but should also be responsible for defining the alternative activities available for meeting the objectives at stake. It is the definition of the options which gives the operating agencies the ability to define the "need" which the standard functions of the agency can meet. It is this ability which has caused the water resources agencies to single-mindedly pursue structural alternatives to the exclusion of a wide range of other non-structural options – for example, congestion charges, flood insurance or floodplain zoning, reduction of water sale restrictions – which could attain the objectives at a lower social cost.

The inequities generated, while less understood, may be fully as serious as the inefficiencies caused by inadequate project evaluations. Because most of the benefits from these projects are tied to the ownership of fixed assets, primarily land, their value becomes incorporated into property and land values and is reflected in the wealth account of the owners of these assets. Because the owners of irrigated and flood plain lands, barge lines, and enterprises benefiting from subsidized transportation costs are not typically poor or even middle income people, the effect of these subsidies is to increase the inequality in the national distribution of wealth holdings.

The incentives for inefficient private behavior is perhaps the most serious characteristic of federal water resource development policy. As is well recognized, floodplain protection through the erection of structures has not been accompanied by a reduction in flood damages. Indeed, even though many billions of dollars have been spent on flood protection projects since 1936, estimates of national flood damages have increased steadily. The current strategy is to confront potential flood damages with the construction of flood protection works which display a favorable damage averted/cost ratio, financed by the general fund of the Treasury. With the floodplain protected, private development which previously appeared inefficient becomes profitable. With the higher level of development, even the reduced stream flows cause damages in excess of what they would have been without flood protection. In effect, because the services of the improved floodplain are made freely available to any development activity, an elaborate series of inefficient activities are stimulated: uneconomic invasion of the floodplain occurs; land values in the plain rise rapidly; alternative measures of damage aversion are not undertaken even though more effective; political pressure for most subsidized control projects is generated; and in the long run, no real flood damages are averted.

The dimensions of this perverse incentive problem have been recognized for nearly a decade now, and a program of mandatory flood insurance with appropriate incentives for efficient behavior has been studied and proposed. Nevertheless, little corrective action has been undertaken. Although flood insurance has been offered to residents in floodplains at a subsidized rate through the 1968 National Flood Insurance Act, few takers have been found. To residents in floodplains, additional protection works financed by taxpayers appear more profitable than the flood insurance subsidy! Examples of other water resources policies with such perverse incentives are numerous.

7.3. The use of litigation in resolving water resource policy disputes

As indicated above, standards and criteria for the evaluation of proposed water resource investments have been established by legislative and by executive action. Increasingly, the application of these rules by the agencies has been challenged in the courts by private groups opposing the construction of individual proposed projects.

Since individual agencies in the U.S. system are left to decide how standards and criteria are to be applied, when third parties believe a rule has been violated, an independent institution – the judiciary – determines procedural rules. These proceedings involve full revelation of factual claims by both parties (including factual claims adverse to one's position) and the resolution of the dispute by the court based upon its understanding of the factual claims and the validity of the data and logic on which the claims rest.

In the water resources area, litigation regarding public investment appraisals has increasingly concerned the extent to which the analysis accompanying a proposed project fulfills the study requirements of the National Environmental Policy Act (NEPA). Much litigation focuses on the benefit–cost evaluation which accompanies every proposed federal project and on the environmental impact study (EIS). This dual focus exists because NEPA, which mandates the EIS, dictates that an evaluation of economic impacts be a part of the EIS.

Although an analysis of the beneficial and detrimental impacts of disputed projects would be expected to help the court to assess the factual conclusions and the accuracy of assertions regarding the effects of these projects, in fact, these government studies appear to have hindered the judicial process. The use of statistics, computer models, probabilities, and mathematics in these government-sponsored studies has appeared sufficiently forbidding to the courts that extreme and undue deference has tended to be given to the government's position. An accepted presumption is that the conclusions of the government-sponsored report rest soundly on the analysis that underlies it, and that this analysis, in turn, is accurate, pertinent, and appropriate. When this presumption is challenged, however, the government representative often obscures facts and relationships which could be expressed in layman's terms by using scientific jargon, computer modeling, and seemingly complete statistical analysis. This tendency to overwhelm the court with volume and technical detail plays on the inability of the courts to determine if the facts in a situation do or do not comply with a general rule if it is unable to understand the factual evidence presented. The result is the loss of the probing skepticism so critical to a rational decision process, and with it the loss of an important mechanism for insuring public decisions which meet efficiency goals [see Carroll et al. (1983)].

7.4. Concluding remarks on public water policy

Pervading our description of water resource policy has been the problem of the separation of the beneficiaries of resource use from those who bear the cost of such use. Because of this separation, distinguishable groups of people are subsidized at the expense of other (typically larger) groups. Moreover, the subsidized are provided both the incentive and the wherewithal to manipulate the political system to maintain the flow of subsidy – whether or not an economic or social function is served and without regard to cost [Gardner (1983)]. As our discussion has revealed, these groups are readily identifiable – landowners along flood protected watercourses, the owners of industrial and commercial firms contributing to waterborne wasteloads, irrigators using publicly developed water supplies, and the owners of barge lines using public waterways.

Improvements in the efficiency and equity of water resources policy are likely to be achievable by a reduction of the large volume of subsidies conferred by existing policy. Elimination of the existing separation of beneficiaries and cost-bearers of policy measures through a comprehensive beneficiary charge policy could yield this improvement. The major components of such a policy would move toward:

- Effluent charges or marketable effluent permits to discourage waste discharges to public watercourses, coordinated with programs governing solid waste and airborne disposals.
- An increase in the price of publicly-produced irrigation water to reflect its supply cost and opportunity cost in alternative uses.
- A national system of mandatory flood control insurance with premiums set equal to expected loss.
- Imposition of long-run marginal cost-based user charges on barge lines using public waterways.
- Supply cost and congestion-related user charges for waterbased recreation areas.

References

Allee, D.J. (1982) "Failure of the Public Market – A Framework for Cost-Sharing Policy Research", in: R.M. North and S.H. Hanke (eds.), *Financing Water Resources: Cost Allocation, Cost Sharing and Incentives* (Institute of Natural Resources, University of Georgia, Athens).

Anderson, R.L. (1961) "The Irrigation Water Rental Market, A Case Study", *Agricultural Economics Research* 8, 54–58.

Anderson, R.L. and A. Maass (1971) *A Simulation of Irrigation Systems*, Tech. Bull. 1431, ERS (U.S. Dept. of Agriculture, Washington).

Anderson, T.L. (ed.) (1983) *Water Rights: Scarce Resource Allocation, Bureaucracy and the Environment* (Ballinger Publishing Co., Cambridge).

Angelides, S. and E. Bardach (1978) *Water Banking: How To Stop Wasting Agricultural Water* (Institute for Contemporary Studies, San Francisco).

Ansari, Nasim (1968) *Economics of Irrigation Rates* (Asia Publishing House, New York).

Arrow, K.J. and R.C. Lind (1970) "Uncertainty and the Evaluation of Public Investment Decisions", *American Economic Review* 60, 364–378.

Baecher, G.B., M.E. Pate and R. De Neufville (1980) "Risk of Dam Failure in Benefit-cost Analysis", *Water Resources Research* 16, 449–456.

Bain, Joe S., Richard E. Caves and Julius Margolis (1966) *Northern California's Water Industry* (Johns Hopkins Press, Baltimore).

Baumgartner, A. and E. Reichel (1975) *The World Water Balance* [English Translation by Richard Lee] (Elsevier, Amsterdam).

Beattie, B.R., E.N. Castle, W.G. Brown and W. Griffin (1971) *Economic Consequences of Interbasin Transfer*, Technical Bulletin 116 (Oregon Agricultural Experiment Station, Corvallis).

Bell, Clive, P. Hazell and R. Slade (1982) *Project Evaluation in Regional Perspective* (Johns Hopkins Press, Baltimore).

Bonbright, James C. (1961) *Principles of Public Utility Rates* (Columbia University Press, New York).

Boulding, K.E. (1964) "The Economist and the Engineer", in: E.N. Castle and S.C. Smith (eds.), *Economics and Public Policy in Water Resource Development* (Iowa State University Press, Ames).

Boulding, K.E. (1980) "The Implications of Improved Water Allocation Policy", in: Marvin Duncan (compiler), *Western Water Resources: Coming Problems and Policy Alternatives* (Westview Press, Boulder).

Bowen, Richard and R.A. Young (1983) *Allocative Efficiency and Equity of Alternative Methods of Charging for Irrigation Water: A Case Study in Egypt*, Egypt Water Use Project Tech. Rept. No. 37, Colorado State University, Fort Collins.

Bower, B.T. (1963) "Some Physical, Technological and Economic Characteristics of Water and Water Resource Systems: Implications for Administration", *Natural Resources Journal* 3, 215–238.

Bradley, I.E. and J.P. Gander (1968) *The Economics of Water Allocation in Utah: An Input–output Analysis*, Bureau of Economic and Business Research (University of Utah, Salt Lake City).

Bredehoeft, J.D. and R.A. Young (1970) "Temporal Allocation of Groundwater: A Simulation Approach", *Water Resources Research* 6, 1–12.

Bromley, D.W., D.C. Taylor and D.E. Parker (1980) "Water Reform and Economic Development: Institutional Aspects of Water Management in the Developing Countries", *Economic Development and Cultural Change* 28, 365–387.

Bromley, D.W. (1982) "Improving Irrigated Agriculture: Institutional Reform and the Small Farmer", Staff Working Paper 531 (World Bank, Washington).

Brown, G. and C.B. McGuire (1967) "A Socially Optimal Pricing Policy for a Public Water Agency, *Water Resources Research* 3, 33–43.

Brown, Lee, B. McDonald, J. Tysseling and C. DuMars (1982) "Water Reallocation, Market Proficiency and Conflicting Social Values", in: G.D. Weatherford (ed.), *Water and Agriculture in the Western U.S.: Conservation, Reallocation and Markets* (Westview Press, Boulder,).

Burt, O.R. (1964) "Optimal Resource Use Over Time with an Application to Ground Water", *Management Science* 11, 80–93.

Burt, O.R. (1975) "Ground Water Management and Surface Water Development for Irrigation", in: R.M. Thrall (ed.), *Economic Modelling for Water Policy Evaluation* (North Holland, Amsterdam).

Butcher, W.R., N.K. Whittlesey and J.F. Osborn (1972) *Economic Values of Water in a System Context*, National Water Commission Technical Report (Nat. Tech. Info. Svc. No. PB210357, Springfield).

Carroll, J., J. Karaganis and R.H. Haveman (1983) "When Complex Facts Threaten Court Reviews: Litigation Over Navigation Projects", *Journal of Policy Analysis and Management* 2, 418–431.

Carruthers, I. and C. Clark (1981) *The Economics of Irrigation* (Liverpool University Press, Liverpool).

Castle, E.N. (1978) "Property Rights and the Political Economy of Resource Scarcity", *American Journal of Agricultural Economics* 60, 1–9.

Castle, E.N., M.M. Kelso, J.B. Stevens and H.H. Stoevener (1981) "Natural Resource Economics, 1946–75", in: L.R. Martin (ed.), *Survey of Agricultural Economics Literature*, vol. III (University of Minnesota Press, Minneapolis).

Chankong, V. and Y.Y. Haimes (1983) *Multiobjective Decision-making: Theory and Methodology* (Elsevier North Holland Publ. Co., New York).

Christensen, D.A., A.S. Morton and E.O. Heady (1981) *The Potential Effect of Increased Water Prices on U.S. Agriculture*, Report 101, Center for Agricultural and Rural Development (Iowa State University, Ames).

Cicchetti, C.J., V.K. Smith and J. Carson (1975) "Economic Analysis of Water Resource Investments and Regional Economic Growth", *Water Resources Research* 11, 1–6.

Ciriacy-Wantrup, S.V. (1952) *Resource Conservation: Economics and Policy* (University of California Press, Berkeley).

Ciriacy-Wantrup, S.V. (1967) "Water Economics: Relations to Law and Policy", in: R.E. Clark (ed.), *Waters and Water Rights*, vol. I (Allen Smith Co., Indianapolis).

Ciriacy-Wantrup, S.V. (1969) "Natural Resources in Economic Growth: The Role of Institutions and Policies", *American Economic Review* 51, 1314–1324.

Ciriacy-Wantrup, S.V. and R.C. Bishop (1975) "Common Property as a Concept in Natural Resources Policy", *Natural Resources Journal* 15, 713–728.

Clark, Robert E. (ed.) (1967) *Waters and Water Rights: A Treatise on the Law of Waters and Related Problems*, vol. I–VII (Allen Smith Co., Indianapolis).

Coase, R. (1971) "The Theory of Public Utility Pricing and its Applications", *Bell Journal of Economics* 1, 113–128.

Cochrane, H.C. (1981) "Flood Loss Simulation", *Natural Resources Forum* 5, 31–68.

Cochrane, H.C. and P.C. Huszar (1983) *Increasing the Economic Efficiency and Affordability of Storm Drainage Projects*, Colorado Water Research Institute Completion Report Series 126 (Colorado State University, Fort Collins).

Cohon, J.L. and D.H. Marks (1975) "Review and Evaluation of Multi-Objective Programming Techniques", *Water Resources Research* 11, 208–220.

Cummings, R.G. (1971) "Optimum Exploitation of Groundwater Reserves with Saltwater Intrusion", *Water Resources Research* 7, 1415–1424.

Cummings, R.G. (1974) "Interbasin Water Transfers: A Case Study in Mexico" (Johns Hopkins Press, Baltimore).

D'Arge, R. (1970) *Quantitative Water Resource Basin Planning: An Analysis of the Pecos River Basin, New Mexico*, Report No. 8 (Water Resources Research Institute, New Mexico State University, Las Cruces).

Daubert, J.T. and R.A. Young (1981) "Recreational Demand for Maintaining Instream Flows", *American Journal of Agricultural Economics* 63, 666–676.

Daubert, J.T., R.A. Young and H.J. Morel-Seytoux (1985) *Managing Interrelated Stream-aquifer Systems; Economics, Institutions and Hydrology* (Colorado Water Resources Research Institute Technical Report 47, Fort Collins, Colorado State University).

Davis, R.K. (1968) *The Range of Choice in Water Management* (Johns Hopkins Press, Baltimore).

Davis, R.K. and S.H. Hanke (1971) *Pricing and Efficiency in Water Resource Management*, National Water Commission Technical Report (National Technical Information Service, No. PB 209083, Springfield).

Ditwiler, D. (1975) "Water Problems and Property Rights; An Economic Perspective", *Natural Resources Journal* 15, 663–680.

Dorfman, R. (1962) "Basic Economic and Technologic Concepts", in: A. Maass, et al. (ed.), *Design of Water Resource Systems* (Harvard University Press, Cambridge) pp. 88–158.

Dudley, N.J., W.F. Musgrave and D.T. Howell (1972) "Irrigation Planning 3; The Best Size of an Irrigation Area for a Reservoir", *Water Resources Research* 8, 7–17.

Easter, K.W. and D.E. Welsch (1983) "Socioeconomic Issues in Irrigation Development and Distribution", Staff Report 83-5, Department of Agricultural and Applied Economics (University of Minnesota, St. Paul).

Eckstein, Otto (1958) *Water Resource Development: The Economics of Project Evaluation* (Harvard

University Press, Cambridge).

Eisel, L.M., G.D. Seinwill and R.M. Wheeler, Jr. (1982) "Improved Principles, Standards and Procedures for Evaluating Federal Water Projects", *Water Resources Research* 18, 203–210.

Feinerman, E. and K.C. Knapp (1983) "Benefits from Groundwater Management: Magnitude, Sensitivity and Distribution", *American Journal of Agricultural Economics* 65, 703–710.

Fisher, A.C. (1973) "Environmental Externalities and the Arrow-Lind Public Investment Theorem", *American Economic Review* 63, 722–725.

Flack, J.E. (1981) "Residential Water Conservation", *Journal of the Water Resources Planning Division* (Am. Soc. Civil Engineers) 107, WR1, 85–95.

Flinn, J.C. and W.F. Musgrave (1967) "Development and Analysis of Input–output Relations for Irrigation Water", *Australian Journal of Agricultural Economics* 11, 1–19.

Flinn, J.C. and J.R. Guise (1970) "An Application of Spatial Equilibrium Analysis to Water Resource Allocation", *Water Resources Research* 6, 2.

Fox, I. (1976) "Institutions for Water Management in a Changing World", *Natural Resources Journal* 16, 743–758.

Frederick, K.D. with J.C. Hanson (1982) *Water for Western Agriculture*, RFF Research Paper (Johns Hopkins University Press, Baltimore).

Freeman, A.M., III (1966) "Adjusted Benefit Cost Ratios for Six Recent Reclamation Projects", *Journal of Farm Economics* 48, 1002–1012.

Freeman, A.M., III (1969) "Project Evaluation and Design with Multiple Objectives", in: The Analysis and Evaluation of Public Expenditures: the PPB System, vol. I (U.S. Congress, Joint Economic Committee) pp. 565–578.

Freeman, A.M., III (1979) *The Benefits of Environmental Improvement* (Johns Hopkins Press, Baltimore).

Fullerton, Herbert and others "Regional Development: An Econometric Study of the Role of Water Development in Effectuating Population and Income Change", Rept. PRRBE089-1 (Utah State Water Research Laboratory, Logan).

Gardner, B.D. (1966) "Effect of Resource Policies on Income Distribution", *Journal of Farm Economics* 48, 1242–1253.

Gardner, B.D. and H.H. Fullerton (1968) "Transfer Restrictions and the Misallocation of Irrigation Water", *American Journal of Agricultural Economics* 50, 556–571.

Gardner, B.D. (1983) "Water Pricing and Rent-Seeking in California Agriculture", in: T.L. Anderson (ed.), *Water Rights* (Ballinger Publishing, Cambridge).

Gardner, R.L. and T.A. Miller (1983) "Price Behavior in the Water Market of Northeastern Colorado", *Water Resources Bulletin* 19, 557–566.

Gittinger, J.P., III (1982) *Economic Analysis of Agricultural Projects*, 2nd ed. (Johns Hopkins Press, Baltimore).

Goicoechea, A., D.R. Hansen and L. Duckstein (1982) *Multiobjective Decision Analysis with Engineering and Business Applications* (Wiley, New York).

Gorelick, S.M. (1983) "A Review of Distributed Parameter Groundwater Management Modeling Methods", *Water Resources Research* 19, 305–319.

Gray, S.L. and R.A. Young (1984) "Valuation of Water on Wildlands", in: G.K. Peterson and A. Randall (eds.), *Valuation of Wildlands Resource Benefits* (Westview Press, Boulder).

Griffin, R.C. and J.R. Stoll (1983) "The Enhanced Role of Water Conservation in Benefit-cost Analysis", *Water Resources Bulletin* 19, 447–458.

Gum, R.L. and others (1982) "Quantified Evaluation for Decisions: Measuring Environmental Goals and Quantifying Tradeoffs", Technical Bulletin 145 (Oregon Agricultural Experiment Station, Corvallis).

Gysi, M. (1980) *Guide to Conservational Pricing* (Vantage Press, New York).

Haimes, Y.Y. (ed.) (1981) *Risk-benefit Analysis in Water Resources Planning* (Plenum, New York).

Hanke, S.H. (1970) "Demand for Water Under Dynamic Conditions", *Water Resources Research* 6, 1253–1261.

Hanke, S.H. and R.K. Davis (1974) "The Role of User Fees and Congestion Tolls in Management of Inland Waterways", *Water Resources Bulletin* 10, 54–65.

Hanke, S.H., R.H. Carver and P. Bugg (1975) "Project Evaluation During Inflation", *Water Resources Research* 11, 511–514.

Hanke, S.H. and R.W. Wentworth (1981) "Project Evaluation During Inflation, Revisited", *Water Resources Research* 17, 1737–1738.

Harberger, A.C. (1968) "On Measuring the Opportunity Costs of Public Funds", in: *Proceedings, Committee on Water Resources and Economic Development of the West: The Discount Rate in Public Investment Evaluation* (Western Agricultural Economics Research Council, Denver).

Hartman, L.M. and D.A. Seastone (1970) *Water Transfers: Economic Efficiency and Alternative Institutions* (Johns Hopkins Press, Baltimore).

Haveman, R.H. (1965) *Water Resource Investments and the Public Interest* (Vanderbilt University Press, Nashville).

Haveman, R.H. (1969) "The Opportunity Cost of Displaced Private Spending and the Social Discount Rate", *Water Resources Research* 5, 947–957.

Haveman, R.H. (1972) *The Economic Performance of Public Investments: An ex post Evaluation of Water Resources Investments* (Johns Hopkins University Press, Baltimore).

Haveman, R.H. (1973) "Common Property, Congestion and Environmental Pollution", *Quarterly Journal of Economics* 88, 58–68.

Haveman, R.H. and P. Stephan (1968) "The Budget Congress won't Cut", *The Reporter* 22, 36–47.

Haveman, R.H. and J.V. Krutilla (1968) *Unemployment, Idle Capacity, and the Evaluation of Public Expenditures* (Johns Hopkins University Press, Baltimore).

Heady, E.O. (1952) *Economics of Agricultural Production and Resource Use* (Prentice Hall, New York).

Heaney, J.P. and R.E. Dickinson (1982) "Methods for Apportioning the Cost of a Water Resource Project", *Water Resources Research* 18, 476–482.

Heiner, R.A. (1983) "The Origin of Unpredictable Behavior", *American Economic Review* 73, 560–595.

Herfindahl, O. and A.V. Kneese (1974) *Economic Theory of Natural Resources* (Merrill Publishing, Columbus).

Hexem, Roger W. and Earl O. Heady (1978) *Water Production Functions for Irrigated Agriculture* (Iowa State University Press, Ames).

Hirshleifer, J., J.C. DeHaven and J.A. Milliman (1969) *Water Supply: Economics, Technology and Policy* (The University of Chicago Press, Chicago).

Holmes, B.H. (1972) *A History of Federal Water Resources Programs, 1800–1959*, Misc. Publ. 1233 (U.S. Dept. of Agriculture, Washington).

Holmes, B.H. (1979) *History of Federal Water Resources Programs and Policies, 1961–1970*, Misc. Publ. 1379 (U.S. Dept. of Agriculture, Washington).

Howe, C.W. (1971) *Benefit-Cost Analysis for Water System Planning*, Water Resources Monograph No. 2 (American Geophysical Union, Washington).

Howe, C.W. (1976) "Effects of Water Resource Development on Economic Growth: the Conditions for Success", *Natural Resources Journal* 6, 939–956.

Howe, C.W. (1979) *Natural Resource Economics: Issues, Analysis and Policy* (Wiley, New York).

Howe, C.W. (1982) "Impact of Price on Residential Water Demand: New Insights", *Water Resources Research* 18, 713–716.

Howe, C.W. and K.W. Easter (1971) *Interbasin Transfers of Water: Economic Issues and Impacts* (Johns Hopkins University Press, Baltimore).

Howe, C.W. with others (1984) *Innovations in Water Management: An ex post Appraisal of the Colorado–Big Thompson Project and the Northern Colorado Water Conservancy District* (monograph in Preparation, University of Colorado).

Howitt, R.E., W.D. Watson and R.M. Adams (1980) "A Reevaluation of Price Elasticities for Irrigation Water", *Water Resources Research* 16, 623–628.

Howitt, R.E., D.E. Mann and H.J. Vaux, Jr. (1982) "The Economics of Water Allocation", in: E.A. Englebert (ed.), *Competition for California Water* (University of California Press, Berkeley).

Huffman, R.E. (1953) *Irrigation Development and Public Water Policy* (Ronald Press, New York).

Huffschmidt, M.M. and M.B. Fiering (1966) *Simulation Techniques for Design of Water Resource Systems* (Harvard University Press, Cambridge).

Huffschmidt, M.M. and J. Gerin (1970) "Systematic Errors in Cost Estimates for Public Investment Projects", in: J. Margolis (ed.), *The Analysis of Public Output* (Columbia University Press, New York).

Infanger, C.L. and W.R. Butcher (1974) "Individual Income Redistribution and Publicly Provided

Irrigation: The Columbia Basin Project", *American Journal of Agricultural Economics* 56, 805–811.

James, L.D. and R.R. Lee (1971) *Economics of Water Resources Planning* (McGraw-Hill, New York).

Joeres, E.F. and M.A. David (eds.) (1983) *Buying a Better Environment: Cost Effective Regulation Through Permit Trading* (University of Wisconsin Press, Madison).

Johnson, G.E. and P.R.G. Layard (1982) "Public Employment with Labor Distortions", in: R.H. Haveman (ed.), *Public Finance and Public Employment* (Wayne State University Press, Detroit).

Johnson, R.N., M. Gisser and M. Werner (1981) "The Definition of a Surface Water Right and Transferability", *Journal of Law and Economics* 24, 273–288.

Johnson, S.H., III (1982) "Large Scale Irrigation and Drainage Schemes in Pakistan", *Food Research Institute Studies* 18, 149–180.

Kates, R.W. (1970) *Natural Hazards in Flood Plain Management*, Paper No. 14, Natural Hazards Research Program (University of Colorado, Boulder).

Kelso, M.M. (1967) "The 'Water is Different' Syndrome: or What is Wrong with the Water Industry?", *Proceedings, Third Annual Conference* (American Water Resources Association) pp. 176–183.

Kelso, M.M., W.E. Martin and L.E. Mack (1973) *Water Supplies and Economic Growth in an Arid Environment: An Arizona Case Study* (University of Arizona Press, Tucson).

Kimball, N.D. and E.N. Castle (1963) *Secondary Benefits and Irrigation Project Planning*, Technical Bulletin No. 69 (Oregon Agricultural Experiment Station, Corvallis).

Kneese, A.V. and C.L. Schultze (1975) *Pollution, Prices and Public Policy* (Brooking Institution, Washington).

Knetsch, J.L. (1974) "Outdoor Recreation and Water Resources Planning", *Water Resources Monograph No. 3* (American Geophysical Union, Washington).

Krutilla, J.V. and O. Eckstein (1958) *Multiple Purpose River Development: Studies in Applied Economic Analysis* (Johns Hopkins Press, Baltimore).

Krutilla, J.V. (1966) "An Economic Approach to Coping with Flood Damage", *Water Resources Research* 2, 186–194.

Kunreuther, Howard (1978) *Disaster Insurance Protection* (Wiley, New York).

Lewis, W.C. and others (1973) *Regional Growth and Water Resource Investment*" (Lexington Books, Lexington).

Lind, R.C. (1967) "Flood Control Alternatives and the Economics of Flood Protection", *Water Resources Research* 3, 345–357.

Lind, R.C. (ed.) (1982) *Discounting for Time and Risk in Energy Policy* (Johns Hopkins University Press, Baltimore).

Loehman, E., J. Orlando, J. Tschirhart and A. Whinston (1979) "Cost Allocation for a Regional Wastewater Treatment System", *Water Resources Research* 15, 193–202.

Lofting, E.M. and P.H. McGauhey (1968) "Economic Evaluation of Water, Part IV: An Input–output and Linear Programming Analysis of California Water Requirements", Contribution No. 116 (Water Resources Center, University of California, Berkeley).

Lord, W.B. (1979) "Conflict in Federal Water Resources Planning", *Water Resources Bulletin* 15, 1235–1266.

Loughlin, J.C. (1977) "The Efficiency and Equity of Cost Allocation Methods for Multipurpose Water Projects", *Water Resources Research* 13, 8–15.

Maass, A. and others (1962) *Design of Water-resource Systems* (Harvard University Press, Cambridge).

Maass, A. and R.L. Anderson (1978) *And the Desert Shall Rejoice: Conflict, Growth and Justice in Arid Environments* (Massachusetts Institute of Technology Press, Cambridge).

Major, D.C. and R.L. Lenton (1979) *Applied Water Resource Systems Planning* (Prentice Hall, Englewood Cliffs).

Mann, D.E. (1982) "Institutional Framework for Agricultural Water Conservation and Reallocation", in: G.D. Weatherford (ed.), *Water and Agriculture in the Western U.S.; Conservation, Reallocation, Markets* (Westview Press, Boulder).

Mapp, H.P. and V.R. Eidman (1976) "A Bioeconomic Simulation Analysis of Regulating Groundwater Irrigation", *American Journal of Agricultural Economics* 58, 391–402.

Marglin, Stephen A. (1962) "Objectives of Water Resources Development: A General Statement", Chapter 2 of Arthur Maass, et al., *Design of Water Resource Systems* (Harvard University Press, Cambridge).

Marglin, S.A. (1963) *Approaches to Dynamic Investment Planning* (North Holland, Amsterdam).

Marglin, S.A. (1968) "The Discount Rate in Public Investment Evaluation", in: *Proceedings, Committee on Economics of Water Resources Development The Discount Rate in Public Investment Evaluation*, Rept. No. 17 (Western Agricultural Economics Research Council, Denver).

Marshall, H.E. (1970) "Economic Efficiency Implications of Federal-local Cost-sharing in Water Resources Development", *Water Resources Research* 6, 673–682.

Marshall, H.E. and V.L. Broussalian (1972) *Federal Cost-sharing Policies for Water Resources*, National Water Commission Technical Report (National Technical Information Service No. PB 208 304, Springfield).

Martin, W.E. (1978) "Economies of Size and the 160-Acre Limitation: Fact and Fancy", *American Journal of Agricultural Economics* 60, 923–934.

Martin, W.E. (1979) "Returns to Public Irrigation Development and the Costs of Commodity Programs", *American Journal of Agricultural Economics* 61, 1107–1114.

McKean, Roland N. (1958), *Efficiency in Government Through Systems Analysis* (John Wiley and Sons, New York).

Millan, J. (1972) *Drought Impact on a Regional Economy* (Unpublished Ph.D. Dissertation, Colorado State University).

Miller, J.R. and D.A. Underwood (1983) "Distributional Issues in Western Municipal and Industrial Water Supply", *Water Resources Bulletin* 19, 631–640.

Milliman, Jerome (1972) "Beneficiary Charges: Toward a Unified Theory", in: Selma Mushkin (ed.), *Public Prices for Public Products* (Urban Institute, Washington) pp. 27–52.

Milliman, J.W. (1983) "An Agenda for Economic Research on Flood Hazard Mitigation", in: S.A. Changnon, Jr., et al. (ed.), *A Plan for Research on Floods and their Mitigation in the United States* (processed, Illinois State Water Survey, Champaign).

Mishan, E.J. (1976) *Cost-benefit Analysis*, 2nd ed. (Praeger, New York).

Moore, C.V. (1961) "Analytical Framework for Estimating the Production Function for Crops using Irrigation Water", *Journal of Farm Economics* 43, 876–888.

Moore, C.V. (1972) "On the Necessary and Sufficient Conditions for a Long-term Irrigated Agriculture", *Water Resources Bulletin* 8, 802–812.

Moore, C.V. (1981) "Economic Evaluation of Irrigation with Saline Water within the Framework of a Farm: Methodology and Empirical Findings", in: D. Yaron (ed.), *Salinity in Irrigation and Water Resources* (Marcel Dekker, Inc., New York).

Moore, C.V. (1982) "Impact of Federal Acreage Limitation Policy on Western Irrigated Agriculture", *Western Journal of Agricultural Economics* 7, 301–315.

Moore, N.Y. and W. Yeh (1980) "Economic Model for Reservoir Planning", *Journal of Water Resources Planning and Management Division, American Society of Civil Engineers* 106(WR-2), 383–400.

Munasinghe, M. and J.J. Warford (1982) *Electricity Pricing: Theory and Case Studies* (Johns Hopkins Press, Baltimore).

Musgrave, W.F. and J.F. Lesueur (1973) "Large Dams and Drought", in: J.V. Lovett (ed.), *The Environmental, Economic and Social Significance of Drought* (Angus and Robertson, Sidney).

Nelson, J.R. (1969) "Policy Analysis in Transportation Programs", in: *The Analysis and Evaluation of Public Expenditures* (Joint Economic Committee, 91st Congress, Washington) pp. 1102–1127.

North, Ronald K. and W.P. Neely (1976) "A Model for Achieving Consistency for Cost-sharing for Water Resource Programs", *Water Resources Bulletin* 13, 995–1005.

Oyarzabal-Tamargo, F. and R.A. Young (1978) "International External Diseconomies: The Colorado River Salinity Problem in Mexico", *Natural Resources Journal* 18, 77–88.

Pearce, D.W. and C.A. Nash (1981) *The Social Appraisal of Projects; A Text in Cost-benefit Analysis* (John Wiley, New York).

Randall, Alan (1981) "Property Entitlements and Pricing Policies for a Maturing Water Economy", *Australian Journal of Agricultural Economics* 25, 195–212.

Randall, Alan (1983) "The Problem of Market Failure", *Natural Resources Journal* 23, 131–148.

Russell, C.S., D.G. Arey and R.W. Kates (1970) "Drought and Water Supply" (Johns Hopkins Press, Baltimore).

Ruttan, V.W. (1965) *The Economic Demand for Irrigated Acreage: New Methodology and Some Preliminary Projections, 1954–1980* (Johns Hopkins University Press, Baltimore).

Ruttan, V.W. (1978) "Induced Institutional Change", in: H. Binswanger and V.W. Ruttan (eds.), *Induced Innovation* (Johns Hopkins Press, Baltimore) pp. 327–356.

Sampath, R.K. (1983) "Returns to Public Irrigation Development; A Theoretical Note", *American Journal of Agricultural Economics* 65, 337–339.

Saunders, R.J., J.J. Warford and P.C. Mann (1977) *Alternative Concepts of Marginal Cost for Public Utility Pricing: Problems of Application in the Water Supply Sector*, Working Paper 259 (World Bank, Washington).

Schmid, A.A. (1972) "Analytical Institutional Economics: Challenging Problems in the Economics of Resources for a New Environment", *American Journal of Agricultural Economics* 54, 893–901.

Schramm, Gunter and F.V. Gonzales (1976) "Pricing Irrigation Water in Mexico, Efficiency, Equity and Revenue Considerations", *Annals of Regional Science* 11, 15–35.

Seagraves, J.A. and K.W. Easter (1983) "Pricing Irrigation Water in Developing Countries", *Water Resources Bulletin* 19, 663–672.

Seckler, D.W. and R.A. Young (1979) "Economic and Policy Implications of the 160-acre Limit in Federal Reclamation Law", *American Journal of Agricultural Economics* 60, 575–588.

Seckler, D.W. and K.C. Nobe (1983) "The Management Factor in Developing Economies", in: K.C. Nobe and R.K. Sampath (eds.), *Issues in Third World Development* (Westview Press, Boulder).

Shabman, Leonard (1982) "Setting Navigation User Fees in Relation to Costs: A Review of the Issues", in: *Proceedings, Waterways User Charges Conference*, Report AE-4535 (Illinois Agricultural Experiment Station, Urbana).

Solley, W.B., E.B. Chase and W.B. Mann (1983) *Estimated Use of Water in the United States*, Circular 1001 (U.S. Geological Survey, Washington).

Sorkin, A.L. (1982) *Economic Aspects of Natural Hazards* (D.C. Heath, Lexington).

Stavins, Robert (1983) "Trading Conservation Investments for Water", Processed (Environmental Defense Fund, Berkeley).

Stockfisch, J.A. (1982) "Measuring the Social Rate of Return on Private Investment", in: R.C. Lind (ed.), *Discounting for Time and Risk in Energy Policy* (Johns Hopkins Press, Baltimore).

Supalla, R.J., et al. (1982) "The Ogallala-High Plains Aquifer Study", *Southwestern Review of Management and Economics* 2, 1–174.

Teele, R.P. (1927) *The Economics of Land Reclamation in the United States* (A.R. Shaw, Chicago).

Trelease, Frank J. (1965) "Policies for Water Law: Property Rights, Economic Forces and Public Regulation", *Natural Resources Journal* 5, 1–48.

U.S. Bureau of the Budget (1952) *Reports and Budget Estimates Relating to Federal Programs and Projects for Conservation, Development and Use of Water and Related Land Resources*, Circular A-47 (Washington).

U.S. Congress (1962) *Policies, Standards and Procedures in the Formulation, Evaluation and Review of Plans for Use and Development of Water and Related Land Resources* (87th Congress, 2nd Session, Senate Document 97).

U.S. Congress (1979) *Issues and Options in Flood Hazard Management* (Office of Technology Assessment, Washington).

U.S. Interagency Committee on Water Resources (1958) *Proposed Practices for Economic Analysis of River Basin Projects*, Revised (The "green book") (U.S. Government Printing Office, Washington).

U.S. National Water Commission (1973) *Water Policies for the Future*, Final Report (U.S. Government Printing Office, Washington,).

U.S. Water Resources Council (1973) "Procedures for Evaluation of Water and Related Land Resources Projects", *Federal Register* 38, 174, 24776–24869.

U.S. Water Resources Council (1978) *The Nation's Water Resources, 1975–2000*, vols. 1–4 (Govt. Printing Office, Washington).

U.S. Water Resources Council (1979) "Procedures for Evaluation of National Economic Development Benefits and Costs in Water Resources Planning", Level C; Final rule; *Federal Register* 44, 242, 72892–72976.

U.S. Water Resources Council (1980) *Essentials of Ground-water Hydrology Pertinent to Water Resources Planning*, Bulletin 16, Revised (U.S. Government Printong Office, Washington).

U.S. Water Resources Council (1983) *Economic and Environmental Principles and Guidelines for Water* *and Related Land Resources Implementation Studies* (Govt. Printing Office, Washington).

Vaux, H.J., Jr. and W.O. Pruitt (1983) "Crop-water Production Functions", in: Daniel Miller (ed.), *Advances in Irrigation* 2, 61–97.

Wade, Robert (1982) "The System of Administrative and Political Corruption: Land Irrigation in South India", *Journal of Development Studies*, 18, 287–299.

Warren, J.P., L.L. Jones, R.D. Lacewell and W.L. Griffin (1975) "External Costs of Land Subsidence in the Houston-Baytown Area", *American Journal of Agricultural Economics* 57, 450–455.

Warrick, R.A. (1975) *Drought Hazard in the United States: A Research Assessment* (Institute of Behavioral Science, University of Colorado, Boulder).

White, G.F. (1964) *Choice of Adjustment to Floods*, Research Paper 93 (Department of Geography, University of Chicago).

White, G.F. (1969) *Strategies for American Water Management* (University of Michigan Press, Ann Arbor).

Whittlesey, N.K. and K.C. Gibbs (1978) "Energy and Irrigation in Washington", *Western Journal of Agricultural Economics* 3, 1–9.

Wolf, C., Jr. (1979) "A Theory of Non-Market Failures", *The Public Interest* 55, 114–133.

Wollman, Nathaniel (1963) *The Value of Water in Alternative Uses* (University of New Mexico Press, Albequerque).

Wong, B.D.C. and J.W. Eheart (1983) "Market Simulations for Irrigation Water Rights; An Irrigation Case Study", *Water Resources Research* 19, 1127–1138.

Yaron, D. (1979) "Economics of Irrigation and the Institutional and Pricing Systems of Water in Israel", *Agricultural Water Management* 2, 203–216.

Yaron, D. and A. Dinar (1982) "Optimal Allocation of Irrigation Water", *American Journal of Agricultural Economics* 64, 681–689.

Young, C.E., R.K. Kinsley and W.E. Sharpe (1983) "Impact on Residential Water Consumption of an Increasing Rate Structure", *Water Resources Bulletin* 19, 81–88.

Young, H.P., N. Okada and T. Hashimoto (1982) "Cost Allocation in Water Resources Development", *Water Resources Research* 18, 463–475.

Young, R.A. (1978) "Economic Analysis and Federal Irrigation Policy: A Reappraisal", *Western Journal of Agricultural Economics* 3, 257–268.

Young, R.A. (1984) "Direct and Regional Economic Impacts of Competition for Irrigation Water", in: E.A. Englebert and A.F. Scheuring (eds.), *Water Scarcity: Impacts on Western Agriculture* (University of California Press, Berkeley).

Young, R.A., S.L. Gray, with R.B. Held and R.S. Mack (1972) *Economic Value of Water: Concepts and Empirical Estimates*, Report to the National Water Commission, NTIS no. PB210356 (National Technical Information Service, Springfield).

Chapter 12

MULTIPLE USE MANAGEMENT OF PUBLIC FORESTLANDS*

MICHAEL D. BOWES and JOHN V. KRUTILLA[†]

Resources for the Future, Washington

1. Introduction

This chapter is intended to provide a background in economic concepts specifically focused on the multiple use management of public forestlands. In doing this we use and substantially extend some of the models introduced in Chapter 2 of this Handbook. The managers of these lands must, in addition to considering the value of timber harvests, the primary focus of Chapter 2, consider the various nonmarket amenity services such as recreation, waterflow and wildlife which are influenced by alterations in the standing stocks of timber. We discuss research results on the relation of such multiple use management to single purpose timber management. The presentation is motivated by a number of issues of current concern: the withdrawal of lands from timber management, the specialization or diversification of land use, the level and stability of timber supply, and the wisdom of certain accepted rule-of-thumb principles of public forestland management related to the age and level of harvests.

Our approach is to consider the optimal scheduling of harvests on a single land area of homogeneous productivity on which there are distributed stands of timber of various ages. In each time period the manager may choose to harvest some fraction of each age class. He does so to maximize the discounted stream of net benefits from the various goods and services. As a result of ecological or visual interdependence among stands in the provision of amenity services, the decision

---

* The research on which this chapter is based was undertaken as part of the Forest Economics and Policy Program of Resources for the Future and is supported by funds from Resources for the Future, the U.S. Department of Agriculture Forest Service, and the Weyerhaeuser Company Foundation.

[†] We are indebted to our former colleagues, William F. Hyde, Anthony C. Fisher, and V. Kerry Smith for comments on an earlier draft of this chapter. An equal debt is also acknowledged to Peter Berk for his detailed comments on an earlier draft. All errors of analysis or judgment are the responsibility of the authors.

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

to harvest a particular stand may reflect a complex balancing of the benefit from timber harvests against the benefits of a desirable age distribution in the standing stock. We take benefits to be demand based, consumer surplus type measures and take these and all aspects of production to be known with certainty. The approach which, of course, represents a great simplification of the actual multiple use planning problem, does allow us to illustrate many of the interesting aspects of economically motivated multiple use management.

The plan of the chapter is as follows. First, in Section 2 there is a brief introduction to the institutional setting for multiple use management within the U.S. Forest Service. In Section 3 there is a presentation of the traditional timber harvest scheduling literature and a simple extension that accounts for amenity services. The focus of traditional models is the harvesting of a single stand. In Section 4 a more general multiple use harvest scheduling problem is described. The distinguishing feature is the interdependence of a set of stands in providing amenity benefits. In Section 5 the nature of the harvest solution, particularly the steady state solution, is described. In Section 6 the various conclusions are illustrated and related to policy concerns using some simple examples.

2. Public forestry in the United States

The principles of public forestry in the United States have often been driven by a fear of timber shortages and the destabilizing effects on employment and prices arising from fluctuations in the supply of timber. A distrust of the private market as a supplier of timber arose in reaction to the rapid clearing of forestlands that came with the settlement of the West and as a response to the thefts of timber from remaining public lands. These public forests were, for all practical purposes, common property, unprotected against trespass until after the Organic Administration Act of 1897.[1] The Organic Act authorized the active management of previously set aside public forest reserves. It is viewed as a compromise between advocates of preservation and private exploitation. The management responsibilities for the public forest reserves were transferred in 1905 to the U.S. Department of Agriculture. The philosophy of management [associated particularly with Gifford Pinchot, the first chief of the USDA Forest Service (1905)] became known as conservationism. It was a philosophy of use based on biological and technocratic, rather than economic, principles. The goal was the promotion of a high perpetual level of services from the public forestlands.

Such concerns for protection, stability, and high level of yield led eventually to a widespread acceptance of some rule-of-thumb principles of public forest man-

---

[1] See USDA Forest Service (1978), a handbook of laws relating to Forest Service activities, for a complete reference to the statues.

agement. Among these principles were that a forest should be harvested so as to move toward a "fully regulated" condition with a sustained and even flow of harvest and that this flow should be at the maximum biological yield. This maximum sustained yield–even flow philosophy has long been criticized as devoid of economic rationale.[2] Indeed it has little to say about the wise investment of scarce budgets among lands of varying productivity. Certainly such a policy does not indicate how one should respond to changing patterns of demands for the various products and amenity services of the forest which, in addition to timber, include water flow, wildlife, range, wilderness and other recreational services which depend upon the condition of the standing vegetation.

While the Forest Service has always had a stated concern for the nontimber resource services, it was not until after the Second World War, with the rapid increase in demand for both timber and outdoor recreation, and in response to political pressures on behalf of such single purposes, that the need for explicitly balanced operating criteria became apparent.[3] Attempts to extend the maximum sustained yield–even flow philosophy to all outputs have, not surprisingly, been ambiguous. Now, under the Multiple Use–Sustained Yield Act of 1960 and the Forest and Rangeland Renewable Resources Planning Act of 1974 (RPA) as amended by the National Forest Management Act of 1976 (NFMA), the Forest Service has been given the legislative mandate to manage the forest so as to maximize the benefits from sustained yield–multiple use production, with consideration given to the relative values of all resources and subject to preserving the productive potential of the land. Despite considerable ambiguity, most would agree that the legislation at least accommodates economic concerns and requires a considerable amount of analysis of economic efficiency [see Haigh and Krutilla (1980)].

In response to the RPA/NFMA legislation the Forest Service has begun an ambitious program of planning for some 190 million acres of National Forest lands by the use of linear programming methods. Large linear programming models [Johnson et al. (1980)] are currently in use or under development on many forests where they are to be used for the evaluation of alternative options for scheduling harvests and land use over time. The planning decision still seems to be heavily constrained by policy. The treatments considered generally reflect prior restrictions on minimum harvest ages as well as standards on the intensity of harvest activity. The harvest age restriction, designed to ensure a high average level of harvest yield, is now often justified as a means of protecting amenity values through long rotation cycles. Standards, often intended to protect amenity

---

[2] See, for example, the papers by Samuelson (1976), Hirshleifer (1976), and Downs (1976).

[3] Zivnuska (1961), Behan (1967), and Alston (1972) provide some history of the multiple use doctrine. Other general discussion of multiple use concerns can be found in Gregory (1955), Walter (1977), and Alston (1979).

values, tend to limit the possibility for advantageous specialization of uses on different areas of the forest. The planning effort has served to highlight the need for clarification of the inherent contradictions between the biological–technocratic principles of which the Forest Service is not entirely free, and those principles of allocative efficiency implied by the balancing of supply and demand. We turn now to a review of the prior literature on timber economics.

3. The forest harvesting literature

There is, indeed, a long tradition of analysis of the economics of timber harvest timing. Since scheduling of the timber harvest is the dominant means for altering the service quality of the land unit it will be helpful to review this literature briefly. Indeed, existing multiple use models in which nontimber values depend on the conditions of single stands are simple generalizations of timber harvest models. This discussion complements and extends the general treatment of bioeconomic models found in Chapter 2 of this Handbook.

3.1. The Faustmann problem

The traditional economic problem is the timing of a sequence of harvests on a stand which is managed in perpetuity with prices, costs, interest rates, and stand productivity assumed known and unchanging. The objective is the maximization of the present value of net receipts. This simple formulation is referred to as the Faustmann model, following an 1849 paper by Martin Faustmann. A large and growing number of writers have described the Faustmann model.[4] Samuelson (1976) gives an excellent discussion that stresses the equivalence of competitive market solutions to the solution of a Faustmann type model.

Beginning with a single unstocked stand, a manager incurs per acre regeneration cost C. The timber grows according to the volume function $V(T)$, giving saleable volume per acre as a function of the chosen harvest age T. At the harvest date, the manager receives revenue $PV(T)$, with price P given net of any harvest costs. The land is regenerated again and the cycle repeats. With prices, costs, the growth function, and the interest rate, i, known and unchanging over time, the selected harvest age should be the same in each subsequent rotation cycle. The problem is to find the rotation length T to maximize the present value of receipts from the current and all future harvest cycles. That is, representing this maximum

---

[4] See Pearse (1967), Clark (1976), Hyde (1980), and also Samuelson (1976) who provides many other references. Gaffney (1967) is a good source within the forestry literature.

present value of the stand by λ, we find age T which solves[5]

$$\lambda = \max_{T} \{ (PV(T)e^{-iT} - C)/(1 - e^{-iT}) \}. \tag{1}$$

The first-order condition for the maximization in problem (1) gives us:

$$PV'(T) - iPV(T) - i\lambda = 0. \tag{2}$$

The interpretation is that the Faustmann harvest should be at the age T_F at which the marginal increase in value from further growth just equals the opportunity cost of delaying the harvest. This opportunity cost includes the potential interest income foregone on the delayed receipt of current harvest revenues plus the interest costs of delaying revenues from future harvest cycles.[6] This later term reflects an implicit rental cost of land. Alternatively, we may describe the solution as indicating that we should hold timber stocks uncut until the rate of growth in the combined asset value of the timber and land itself just equals the market rate of interest. That is, we do not harvest until $PV'(T)/(PV(T) + \lambda) = i$. The land value λ is determined so that the investment in timber management will provide no greater or lesser return than is available from other assets.

No difficulty is introduced if we suppose the initial stand to be stocked with growing timber. The harvest age solution is unaffected, unless, of course, the current age exceeds the Faustmann age. If the stand exceeds the Faustmann age, it is to be harvested immediately and subsequently should follow the Faustmann rotation.

There is the possibility that the greatest present value of land under any timber rotation is negative. Under such conditions we would be better leaving the land unmanaged. If the land were currently stocked and an initial harvest were profitable, we would under these circumstances, cut once and then abandon management. Consideration of nontimber values might, of course, make this decision to leave lands unstocked a less attractive solution, as we shall see later.

Comparative statics of the Faustmann solution

To illustrate some characteristics of the Faustmann solution, it is convenient to rewrite eq. (2) by substituting explicitly for the land value λ. We then find an alternative version of the Faustmann condition:

$$PV'(T)/(PV(T) - C) = i/(1 - e^{-iT}). \tag{2'}$$

---

[5] Note $1/(1 - e^{-iT})$ is equal to the infinite sum $1 + e^{-iT} + (e^{-iT})^2 + \cdots$.

[6] The characteristics of timber growth are such that we would usually meet the second-order condition $V''(T) - iV'(T) < 0$. We may assume that there is some τ such that for $T < \tau$ we have $V(T) = 0$, while for $T > \tau$ we have $V'(T) > 0$ and $V''(T) < 0$.

The value of $i/(1-e^{-iT})$, the right-hand side of eq. (2′) is a downward-sloping function of the rotation age T. As T becomes large its value approaches the interest rate from above. This function is drawn in Figure 12.1 for an interest rate of 4 percent. The left-hand side of eq. (2′) can be interpreted as the relative growth rate in net harvest time revenues. For illustrative purposes we have graphed such a relative growth rate based on Douglas fir yields. The harvest age solution is the age T_F at which the two curves intersect. The solution calls for a relative growth rate in harvest value slightly in excess of the interest rate. Second-order conditions indicate that for the intersection to correspond to a maximum, rather than a minimum, the relative growth rate curve must cut through curve $i/(1-e^{-iT})$ from the above left, as illustrated.

The comparative static results are easily illustrated with the aid of Figure 12.1. With positive regeneration costs, a higher constant price level leads to a shorter rotation age solution. To see this, note that a higher price level shifts the relative timber value growth curve downward. The result is an intersection with curve $i/(1-e^{-iT})$ at a lower solution age. Similarly, a higher cost level shifts the relative value growth rate curve upward, leading to a longer rotation. A higher interest rate raises the values $i/(1-e^{-iT})$ and leads to a shorter rotation age. These results are easily shown by the usual differentiation of the first-order necessary condition. With no cost, the solution age [call it T_0 at which $V'(T)/V(T)=i/(1-e^{-iT})$] is independent of the price level. With positive costs, the Faustmann rotation age is above T_0 and approaches this limit as prices rise relative to cost. The greatest possible Faustmann age is that at which

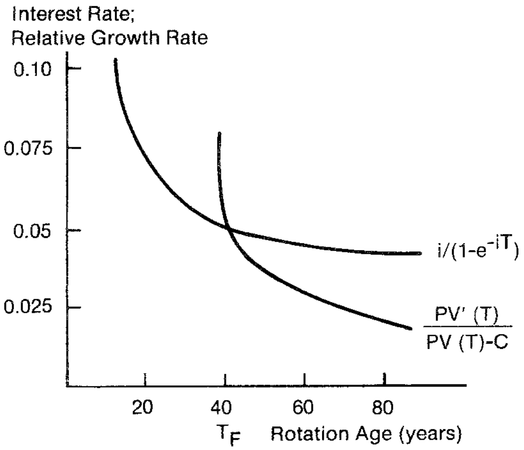

Figure 12.1. The Faustmann rotation age.

$V'(T)/V(T) = i$. This upper bound reflects the limit of profitability where high costs or low price results in a land value just equal to zero.

3.2. *The Hartman problem – multiple use harvesting*

The conventional timber models do not reflect the value of services provided directly by the standing forest. The lack of adequate market signals for nonpriced forest outputs is the heart of the multiple use planning problem and presumably the justification for public management. Hartman (1976) reacted to Samuelson's (1976) discussion with a generalization of the Faustmann harvest problem. His analysis considers one single-aged stand and assumes that the value from nontimber services can be expressed as a function of stand age.

To the basic Faustmann timber problem, Hartman adds a flow of nontimber service values related to the age of the standing stock. The benefit flow from an acre of standing stock of age n is represented by $a(n)$. The integral $\int_0^T (a(n)\mathrm{e}^{-in})\,\mathrm{d}n$ then represents the present value of these amenity services from a single harvest cycle of length T. Beginning with a single unstocked acre, the problem is to choose the harvest age T that will maximize the combined present value of timber and nontimber benefits from the current and future harvest cycles. Again representing this maximized value by λ the problem is given as

$$\lambda = \max_T \left\{ \left[PV(T)\mathrm{e}^{-iT} + \int_0^T (a(n)\mathrm{e}^{-in})\,\mathrm{d}n - C \right] \Big/ (1 - \mathrm{e}^{-iT}) \right\}, \tag{3}$$

with the net price P, volume $V(T)$, and cost C as previously defined for problem (1).[7]

The Hartman rotation age solution T_{H} must satisfy the first-order condition:

$$PV'(T) + a(T) - iPV(T) - i\lambda = 0. \tag{4}$$

This harvest condition (4) is very similar to the Faustmann condition (2) and is again interpreted as calling for a harvest at the age at which the marginal benefits of delaying the harvest just equal the opportunity costs. The benefits of delay include the increment in value from timber growth, plus $a(T)$, the flow of amenity benefits during the period of delay. The costs include the interest income forgone on delayed receipt of the harvest plus the interest cost of delaying the benefits from future harvest cycles. This later term can be interpreted as a rental cost of using land. It now reflects the value of the land for both timber and

---

[7] Compare problem (3) to the similar problem described in Henderson and Quandt (1971, pp. 328–329) on the replacement date for a machine. There, scrap value and product flow corresponds, respectively, to our harvest revenue and amenity flow.

amenity services. We harvest when the rate of increase in the total value of the land and its stock has just fallen to equal the discount rate i.

Because of the varied mix of services they may represent, there is no *a priori* reason to expect the nonharvest values $a(n)$ to be monotonically increasing or decreasing with stand age. It is quite possible that we may find several ages meeting the first-order conditions, some of which are local minima or maxima.[8] With $a(n)$ sufficiently large and increasing with stand age we may find no age at which the first-order conditions are met and would choose to never harvest.

If the maximum land value λ is less than zero then we should leave the land unmanaged. Consideration of the full mix of multiple use values makes the likelihood of such negative returns to management less than if we considered timber values alone. Not surprisingly, we will be less willing to harvest and abandon an area if it provides positive amenity services when stocked. What is interesting is that the amenity values may be sufficient to justify perpetual harvest management on land which would appear to be not economically managed when timber alone is considered. In fact, recreational values, supposedly diminished by timber management, may justify restocking, yet be neither large enough nor increasing sufficiently with stand age as to justify shifting from a typically short timber harvest rotation. This illustrates the danger of an arbitrary allocation of fully joint regeneration costs to specific services.

In contrast to the Faustmann problem, the age of inherited stocks may matter. The harvest age solution is unaffected unless the current age exceeds the Hartman age. If the stand does exceed that age, it may be preferable to further delay, or never, harvest. With timber value alone, the solution was to harvest such older stands immediately. The declining timber growth rate penalized any further delay. Here, with a young stand, discounting may weigh against us waiting for higher amenity values from old growth, yet, if we inherit old growth by chance, the current high flow of amenity value may be sufficient to justify preservation.

In general, the Hartman rotation age will be somewhere between the Faustmann age and that age which would maximize the present value of the returns from the stock amenity services alone. The solution age will depend both on the total amenity benefits of a harvest cycle relative to the net timber receipts and on the separate relative growth rates in the amenity and timber values. When the amenity value of stocks generally rises with stand age, the harvest solution will be greater than the Faustmann age. In fact, if the amenity value flow is large and increasing with stand age, it may be optimal to leave the stand unharvested forever. In many areas where forage and increased water flow are important we might anticipate the amenity values to be declining with stand age. Amenity values generally declining with stand age will lead us to choose shorter rotations.

---

[8] The second-order condition requires $V''(T) - iV'(T) + a'(T)/P < 0$. With $a'(T)$ significant and positive, the condition may be violated.

In order to illustrate these conclusions it proves convenient to rewrite eq. (4). We substitute for the land value λ using definition (3) and then rearrange to decompose the equation into two separate relative growth rates, one for timber and one for amenity services, each weighted by its share of total harvest time benefits. That is, we can equivalently express harvest condition (4) as

$$(1-\alpha)\left[a(T)\Big/\int_0^T\left(a(n)\mathrm{e}^{-in}\right)\mathrm{d}n\right]+\alpha\left[V'(T)/(PV(T)-C)\right]$$
$$=i/(1-\mathrm{e}^{-iT}), \tag{4$'$}$$

where

$$\alpha=[PV(T)-C]\Big/\left[PV(T)+\int_0^T\left(a(n)\mathrm{e}^{-in}\right)\mathrm{d}n-C\right].$$

We may note immediately that with no amenity value, and so with the timber share α equal to one, condition (4′) is simply the Faustmann condition (2′). With high amenity levels and α approaching zero, the harvest solution moves away from the Faustmann age toward the solution age that would be optimal if there were no net timber benefits.[9]

If amenity values are rising with stand age, then the Hartman rotation age exceeds the Faustmann rotation. To see this, note first, using condition (4′), that if the relative growth rate in amenity value $a(T)/\int_0^T(a(n)\mathrm{e}^{-in})\mathrm{d}n$ exceeds the value $i/(1-\mathrm{e}^{-iT})$ for all ages T, then the harvest age solution must exceed the Faustmann timber age. Now, if amenity values rise with stand age, then certainly it is true that the value $\int_0^T(a(T)\mathrm{e}^{-in})\mathrm{d}n$ is greater than $\int_0^T(a(n)\mathrm{e}^{-in})\mathrm{d}n$. On simplification of the first integral, this gives us $a(T)(1-\mathrm{e}^{-iT})/i > \int_0^T(a(n)\mathrm{e}^{-in})\mathrm{d}n$, which is easily rearranged to indicate that the relative growth rate in amenity value does exceed $i/(1-\mathrm{e}^{-iT})$ and so, a longer rotation is called for. Similarly, if amenity values are generally declining with stand age, a rotation shorter than the Faustmann age should be selected.

It is quite possible when we have fairly high amenity values which are generally increasing with stand age that the left-hand side of condition (4′) would always exceed the value $i/(1-\mathrm{e}^{-iT})$. In such a case, the timber should not be harvested. Finally, we note that with a high enough relative growth rate in amenity values, the Hartman multiple use rotation age may significantly exceed the Faustmann age. This may be so even when the share of total benefits due to amenities is rather small in comparison to timber revenues. On the other hand, we may note that if the amenity values are constant with respect to stand age, then, no matter how high these amenity values, the Faustmann timber rotation will be the appropriate harvest age solution. Unfortunately our knowledge tends to be weak

---

[9] It is convenient, although not essential, to assume that $PV(T)-C$ is nonnegative, and so takes values between zero and one.

with respect to the stand-age-dependent rate of change in amenity values. It is apparent that a focus only on the relative level of total benefits, attributable to either timber or other services, can be misleading. The difficulty is compounded once we realize that the value of a stand is likely to be dependent upon the treatment of neighboring areas of the forest.

There are very few examples of papers in which optimal multiple use rotations have been calculated. This is hardly surprising given the difficulty in valuing the amenity services. Calish, Fight and Teeguarden (1978) provided rough estimates of multiple use values for single Douglas fir stands as functions of stand age and solved for optimal rotation ages. Their conclusion was the consideration of these values would have little effect on the harvest solution. This conclusion is perhaps a little too strong.[10] They also note that many of the multiple use services are favored by short harvest rotations. Riiters, Brodie and Hann (1982) in a rather richer model consider the choice of thinning intensity and rotation age on a single stand managed for combined livestock grazing and timber harvest. The consideration of grazing values tends to shorten rotations and leads to intensified thinnings.

Comparative statics of the Hartman solution

The comparative static results are somewhat more complicated for the Hartman solution than for the Faustmann solution. A higher timber price level can be shown to decrease (increase) the rotation age if the Hartman solution T_H is above (less than) age T_0, where T_0 is the Faustmann solution when there are no regeneration costs. An equal proportionate increase in $a(n)$ for each age n lengthens (shortens) the rotation if the Hartman rotation is longer (shorter) than the Faustmann rotation. An increase in all $a(n)$ skewed to a greater increase in the value of older stands would increase the relative rate of growth in the amenity value as well as the amenity share of total benefits and may increase the rotation age even if the initial Hartman solution is somewhat less than the Faustmann age. We will be interested in the effect of an increase in both the timber price and all $a(n)$. For a given proportionate increase in all these timber and amenity prices we can be sure the optimal rotation interval will decrease. More generally, we cannot be sure of the direction of change. In particular, if the increase in values $a(n)$ is skewed to greater increase in value for the older age stands the rotation interval may increase.

It is a little surprising to find that an increase in the interest rate does not necessarily lead to a shorter rotation. With timber value alone considered, the higher rate raises the opportunity cost of delaying the harvest and a shorter

---

[10] They found that large, equal proportionate changes in the amenity values did not greatly alter their solution age. With the low growth rate in amenity value with respect to age in their example, such a result is to be expected. With higher growth rates a more pronounced effect is observed.

rotation results. Here we have two potentially offsetting effects. The value $i/(1-e^{-iT})$ increases with the higher interest rate. This is the same effect we find for the timber solution. However, the greater rate of discounting also raises the relative growth rate in the amenity value, $a(T)/\int_0^T(a(n)e^{-in})\,dn$, decreasing the present value in the denominator. If the share of the amenity benefit is fairly high relative to the timber share, a higher interest rate can lead to longer rotation ages.

3.3. *Maximum sustained yield*

A maximum sustained yield policy is based on the assumption of an even flow of harvest and is aimed at producing the greatest average flow of value per year. Samuelson (1976) compares the Faustmann economic harvest solution to a policy of maximizing average yearly net yield. Such a policy results in the selection of longer rotations and amounts to ignoring interest rates.

The policy can be expressed as the selection of an age T to solve the problem

$$\max_T (PV(T)-C)/T.$$

The first-order condition calls for a harvest age T_M for which we have

$$PV'(T)/(PV(T)-C)=1/T. \tag{5}$$

At the harvest age T_M the increment in growth just equals the average yearly net harvest value. The solution is often referred to in the forestry literature as the "culmination of mean annual increment".

Age T_M must exceed the economic (Faustmann) timber rotation age. To see this, compare expression (5) to the Faustmann condition (2′). It differs only in the right-hand side. It can be easily shown that the value $1/T$ is less than the value of $i/(1-e^{-iT})$, the corresponding term in the Faustmann condition. With the relative growth rate of timber value declining with age, it is apparent that a longer rotation must be selected under this yield-maximizing harvest policy.

The conclusion that the maximum yield policy amounts to ignoring interest rates may be understood by noting that the value $i/(1-e^{-iT})$ approaches $1/T$ as a limit as the interest rate approaches zero. With interest rates close to zero, the Faustmann solution is very close to the yield-maximizing solution. While a policy of higher average yields may seem appealing, one must remember that the timing of the harvest is delayed. This delay imposes costs on the earlier generation of timber users.

It is occasionally suggested that the longer maximum yield rotations are desirable as a means of promoting other multiple use values. Or, it is suggested that the maximum yield concept should be extended to include those other values [Brown and Carder (1977)]. For example, we might consider maximizing the average annual flow of net value from all uses

$$\max_T \left\{ \left[PV(T) + \int_0^T (a(n))\,\mathrm{d}n - C \right] \Big/ T \right\},$$

with the harvest age satisfying the first-order condition

$$[PV'(T) + a(T)] \Big/ \left[PV(T) + \int_0^T (a(n))\,\mathrm{d}n - C \right] = 1/T.$$

This solution corresponds to the Hartman multiple use rotation with a discounting rate of zero. The yield-maximizing harvest ages will only coincidentally equal the economic multiple use rotation.

4. Multiple use management

Johnson and Scheurman (1977) describe linear programming models for timber harvest scheduling which are practical and fairly realistic analogs to the Faustmann model. The FORPLAN model [Johnson, Jones and Kent (1980)] used for the current planning effort of the U.S. Forest Service is an extension of these stand scheduling models to reflect multiple use services. It generalizes, to some extent, the Hartman model. Timber prices and amenity values, although given constants in each time period, may be changing over time. However, even this model, although it schedules many stands at once, has no inherent linkage between stands.

The essential weakness of the Hartman type multiple use models is their focus on the condition of the single timber stand. As Hartman (1976) himself states, when there are "many plots of forest land which could reasonably be taken as units for making cutting decisions, what happens on one plot will clearly affect the value of a standing forest on other units". In this section we consider a model which takes into account this interdependence. The interdependence among stands leads to richer and more appealing harvest decisions than we find in the single stand models.[11]

The multiple use forest production problem can be viewed as the choice among a set of activities which transform the forest stock while providing flows of timber commodities. Both the harvest flow and the services of the standing stock are valued. The manager is to respond to demands in choosing the time sequence of land use treatments. We assume he does so to maximize the present value of consumer plus producer surplus.

---

[11] Nguyen (1979) reconsidered the problem addressed by Hartman. His approach is to constrain the harvest in order that sufficient standing stock remains to meet some level of environmental service.

In order to simplify the discussion, we will accept a less than general description of the planning problem. Under this simplified view, the forest area is assumed to be made of a number of stands each of homogeneous productivity and each of which may be in a different age class. We may decide upon the age at which each stand is to be cut.[12] We consider there to be two classes of demands, that for harvested timber flow and that for amenity services related to the condition of the land and its standing stocks. This latter consumer group represents a diverse set of users of the recreational, wildlife, forage, and water flow services and may include nonusers who value the condition of the forest. The cutting decision determines the current timber harvest and the future progression of the forest condition. The amenity values will depend upon this forest condition, which we take to be a function of the mix of stand ages. In this manner, the amenity outputs can be related to an overall pattern of diversity resulting from the treatment of the individual stands. It would seem most unrealistic to assume that a single stand could be considered independently of the condition of adjacent stands. While it is clear that much of the land's amenity value is dependent on inherent natural features, we focus on the manager's ability to influence the flow of goods and services.

We stress two essential differences between our description of the planning problem and the view that is commonly presented. First, the direct forest outputs are the timber harvest and a measure of the current forest stock condition. This is more than just a matter of practicality, saving us from enumerating the potentially great variety of amenity services of the forest. It allows us to value both qualitative as well as quantitative changes in the final uses. The forest does not directly produce recreation or animal unit months of domestic livestock grazing, but rather produces conditions which may be differently valued because of varying suitability for such final uses. Second, the stock descriptors are taken to be measurable physical characteristics which would in general be based on the interrelation of a number of adjacent stands.

A more general presentation of the planning problem would extend the set of treatment options and a correspondingly richer description of the condition of the forest could then be used. Among the treatments actually available to a manager are a set of possibilities for vegetation manipulation, access restrictions, and trail and facility construction. Options for vegetation manipulation on timberlands may include the timings of thinnings and harvests as well as choices among types of harvesting and regeneration practices. Typically each treatment sequence will alter the services of the whole land area for many uses. It is probably correct that

---

[12] Berck (1981) also discusses the optimal management of renewable resources when *in situ* stocks provide value. He describes a homogeneous stock of timber, its growth rate is related to total volume, and a harvest is the extraction of an unspecified component of the stock. Such a stylized view is not completely suited to the timber problem for which it is easy to identify growth rates and harvest rules specific to each separate age class of timber.

the choice among the fuller set of treatments would indicate much more flexibility in adapting to multiple use demand than is apparent when we consider simple manipulation of the harvest age. But, such generality, although needed in actual practice, comes at great expense in notational clarity. Our simpler view is sufficiently rich for expository purposes and allows us to fit our discussion into the traditional framework of the forest economics literature with its focus upon harvest timing.

In Section 4.1 we describe a model of the forest production system. Section 4.2 then describes the demands for forest commodities and stock services and provides measures of value based on these demands. Finally, in Section 4.3 the management problem is described.

4.1. *Forest production*

The stocks of vegetation held at the beginning of time t are represented by a vector x_t with element $x_t(j)$ giving the acreage of land with stock aged j. At the end of each period we decide whether stocks are to be harvested or left standing. During the period stocks age so that stock $x_t(j)$ will be aged $j+1$ at the time of the harvest decision. The holding of unstocked land is represented by $x_t(0)$. In the case of unstocked lands we may decide whether to maintain the land unstocked or to allow growth to begin.[13]

We represent the simple set of production activities by the vectors h_t and g_t. For stocked lands, the vector element $h_t(j)$ gives the acreage of stock $x_t(j-1)$ harvested, at age j, at the end of time t. The element $g_t(j)$ is the acreage grown to age j left unharvested at the end of t. We assume there is no mortality. For unstocked land $x_t(0)$, a harvest $h_t(1)$ is interpreted as the maintenance of the unstocked condition – whether by clearing or through forgoing regeneration. This gives no harvest yield but avoids any regeneration expense. Regeneration of unstocked land $g_t(1)$ may require some expense – assumed to occur at the beginning of the period t.[14]

We can represent the production system by three sets of equations, the first describes the allocations of stocks among activities, the second describes the stock progression over time, and the third defines harvests and costs.

---

[13] It might be better to separately track newly harvested lands from that left unplanted for some time since the regeneration costs are likely to differ. We do not do so.

[14] In many cases costless natural regeneration may be acceptable while delaying regeneration might require some clearing expense. This can be reflected in the activity costs.

4.1.1. Stock use

The current stocks can be allocated among activities according to a system of $t+N$ equations

$$x_t(j-1)=h_t(j)+g_t(j) \qquad (j=1,\ldots,t+N), \tag{6}$$

with N the age of the oldest stock held at time 1. That is, stocks can be either harvested or left standing after the period of growth. We could easily generalize the stock vector to represent acreage of a given age, under a particular management regime and then allow choice among alternative harvest and regeneration methods.

4.1.2. Stock progression

The stock progression is the result of the current choice of harvest, regeneration, and growth according to the equations

$$\begin{aligned} x_{t+1}(0) &= \sum_{j=1}^{t+N} h_t(j), \\ x_{t+1}(j) &= g_t(j) \qquad (j=1,\ldots,t+N). \end{aligned} \tag{7}$$

That is, $x_{t+1}(0)$, the unstocked land acreage at time $t+1$, is the result of harvests at the end of period t. Land stocked with trees aged j to begin time $t+1$ is the result of forgoing harvest of trees reaching age j at the end of time t.

4.1.3. Harvests and costs

The volume of harvest at the end of time t is the sum of the volumes harvested from each age class:

$$H_t=\sum_{j=1}^{t+N} V(j)h_t(j), \tag{8}$$

where $V(j)$ is the merchantable volume per acre from stands ending the period at age j.[15] The volume $V(1)$ from clearing unregenerated lands is zero. The total cost

---

[15] It is probably better to treat harvests from certain subsets of age classes as distinct products. We can deal with this concern in a rough manner by having a quality-adjusted measure of volume, but this can fail if there are shifts in processing technology or in the final product mix demanded. Furthermore, relating volume to stand age is a simplification that depends on our assumed knowledge of the timing of thinnings and other growth improvement activities and ignores the inherent uncertainties of growth.

at time t is given by

$$C_t = Cg_t(1), \tag{9}$$

where C is the per acre regeneration or management cost. These costs occur at the beginning of period t. We could easily include a cost for maintaining a cleared area, or for any of the other activities. Harvest costs are presumed borne by the purchaser and are reflected in the demand for timber. For convenience we do not make the costs time dependent nor do we consider costs specific to the amenity services.

Certainly more sophisticated models of forest dynamics and costs can be given but our simple version is sufficient for illustrative purposes. A more detailed set of stocks and greater complexity in the activities and costs is easily expressed. While we have concentrated on forestlands, a similar problem, but a rather different set of actions will face the manager of grasslands. Also, we have not focused on the stocks of wildlife presuming that the valuation of the wildlife services can be related reasonably well to the condition of the vegetative habitat (taking userestrictions, such as the hunting regulations, as given).

4.2. The forest demands and valuation

We focus on two broad sets of demands, the demand for stumpage and the demands for the characteristics of the land and its vegetation. We will accept consumer's surplus type measures as convenient welfare indicators.[16] The demands are assumed to be unchanging over time and known with certainty.

4.2.1. Timber harvest

The aggregate demand for harvests from our forest unit is a derived demand for logs as input into the production of final consumer goods. We will express the demand price for an incremental unit of stumpage harvest at time t as a function of H_t, the total harvest offered from our unit:

$$P_t = p(H_t).$$

This function will also depend upon the prices of other inputs and the aggregate levels of harvest from other areas. For convenience we will not explicitly include these variables.[17] Harvest costs are presumed to be incurred by the purchaser and so are reflected in the demand price.

---

[16] See Willig (1976) and McKenzie and Pearce (1982) for a discussion of welfare measures.

[17] Harvests enter this demand function both directly, for given prices in the final products markets, and indirectly through the influence of the harvest on the final product prices. It can be called a reduced-form or general-equilibrium demand curve.

The area under the stumpage demand function, up to the harvest quantity supplied, represented by the benefit function B_t,

$$B(H_t) = \int_0^{H_t} p(H)\,\mathrm{d}H, \tag{10}$$

provides an economic willingness-to-pay measure of value. The marginal benefit to the timber sector of an increment in harvest volume, the derivative of eq. (10), is the net price of stumpage:

$$\mathrm{d}B_t(H_t)/\mathrm{d}H_t = p(H_t) = P_t.$$

It is sometimes mistakenly felt that the focus on the demands of the immediate purchaser ignores some economic value in the final product markets. However, if we accept the assumption of a competitive timber-processing sector, we can be sure that our measure (10) is equivalent to measuring the sum, throughout the timber sector, of all changes in profits and consumer willingness to pay that result from harvesting our unit.[18]

It may be that our site is sufficiently small so that harvests will have no effect on price. When the demand price does not depend on our harvest level then the measure of value in eq. (10) reduces to the market revenue from stumpage sales. The price may still depend upon the aggregate supply from other areas and on final products demand and so not be immediately available.

4.2.2. Demands for the services of standing stock

The standing stocks are valued for their indirect service as inputs into the production of final commodities and services. Holders of grazing rights will value forage levels for the role in livestock production. Recreationists can be viewed as combining the scenic qualities and facilities of the site, along with their time and travel, as inputs into the production of recreation. What is notable is first, that the same standing stock may have value to many people, and second, that the characteristics of stock condition are not marketed commodities.[19] Finding their value can be a formidable statistical task. Typically the demand for the site characteristics will not be directly revealed in markets. However, the value of changed characteristics may be revealed through a shift in either the demand for a related market input or the supply of a related market product.[20] Cross-sectional

---

[18] See Carlton (1979) and Jacobsen (1979).

[19] The stock is much like a public good. Also note that the stock may be valued by nonusers. Measurement of value may be particularly difficult for that group, but our notation is sufficiently general to include them.

[20] See Mäler (1971) or Bradford and Hildebrand (1977).

observations on site use related to site characteristics may allow one to estimate the value of changes at a particular site.

We will express the total valuation of the amenity services of our site, summed over all those individuals who value the site, by

$$A_t = A(x_t). \tag{11}$$

The value in period t depends on the full stock vector x_t held during the period. We assume that pricing is not used as a management control. With no difficulty it could be added as an argument of the value function A_t.

As an example, consider the case of recreation. The individual's demand price for an additional visit to the site can be expressed as a function of m_i, the number of visits the ith individual takes, and $Q(x)$, with Q a vector of descriptors of site condition which may be influenced by the mix of standing stocks, x. With each individual's marginal willingness to pay for a trip given by the demand function $R(m, Q(x))$, we may measure recreationists' willingness to pay for use of the site by the area under the demand function at each time t:

$$A(x) = \sum_i \int_0^{m_i} R(m, Q(x))\,\mathrm{d}m.$$

This expression sums consumer's surplus (plus any actual payments for site entry) over each visitor.

Now, of course, not all such benefits are attributable to management since surely the site would have value even with no management. All that should ever be of interest is the increment in site benefits due a management action affecting stocks x. Of particular interest is the marginal value of a change in the amount of standing stock. In order to make our notation analogous to the earlier Hartman model, we represent by $a_t(n)$ the value of a marginal increase in the area of stock growing to age n during time period t. That is, we define

$$a_t(n) = \partial A(x_t)/\partial x_t(n-1).$$

This is an aggregate, summing the marginal values of all those who value the amenity services of the standing stock. This marginal value is, in general, a function of x_t, the full mix of stocks on the unit, not just a function of the single stand age. The dependence of the multiple use values on the mix of stands makes the harvesting decision much richer than that for the earlier Hartman problem.

4.3. *The multiple use management problem*

Represent by $J(x_t)$ the maximum net present value of the future flow of harvests and stock amenity services from land with stock vector x_t on hand at time t. The full management problem is to find the value $J(x_0)$, choosing the harvest sequence to maximize the net present value of the flow of goods and services attainable from the initial inventory of stocks, x_0. The harvest choice is constrained by the forest production possibilities described by eqs. (6)–(9).

The maximization problem may be represented in a convenient manner by noting that a recursive relationship exists at each time t. That is, we have

$$J(x_t) = \max_{h_t, g_t \geq 0} \{[A(x_t) + B(H_t)]e^{-i} - C_t + J(x_{t+1})e^{-i}\}, \tag{12}$$

with the maximization subject to the production constraints (6) and (7), and with harvest H_t and cost C_t as defined by eqs. (8) and (9). The discounting reflects the treatment of harvest and amenity value as if they were received at the end of the time period while regeneration costs are incurred at the beginning of a period.

We may solve the full management problem in a sequential manner by solving the recursively related subproblems (12) for each time period. The recurrence relation (12), typical of dynamic programming, indicates that $J(x_t)$, the maximized value of the flow of goods and services from stocks held at time t, can be decomposed into two parts, a current period flow of net value and the present value of the future flows from the land in its subsequent stock condition. The overall problem is reduced to selecting the best action for each single period, given the current stock, with consideration for the impact of the current action on the value of the stocks available for the next period. Typically the actual solution will require an iterative search procedure,[21] although for some particular forms of the problem the solution is very simply found using dynamic programming methods (for example, a linear objective function, or when x is limited to a small set of possible values).

4.3.1. Necessary conditions for the maximization

The Kuhn–Tucker conditions for problem (12) can be expressed in a concise form by noting that for each existing age class we must choose a positive level for at least one action.[22] We must either cut or grow, choosing the action with the greatest marginal return. The present value of a marginal increase in the chosen actions must be just zero while the return to a marginal increase in the activity not chosen in the optimal solution must be less than zero. For each time t, for

---

[21] See the optimal control methods described in McDonough and Park (1975) and Lyon and Sedjo (1981) or the gradient-free search method described in Berck (1980). These solve the similar timber harvest problem. The nontimber value function may introduce nonconvexities making solutions more difficult to find than is the case for the corresponding timber problem.

[22] The Kuhn–Tucker necessary conditions require that at the optimum any activity y (either h or g) satisfy: $\partial L/\partial y \leq 0$, $y\,\partial L/\partial y = 0$, and $y \geq 0$. Suppose $\partial L/\partial y < 0$, then we must have $y = 0$ (that is, the action cannot be used). Similarly, if we use an action with $y > 0$ then we can meet the required condition $y\,\partial L/\partial y = 0$ only if $\partial L/\partial y = 0$.

Table 12.1
Summary and definitions for terms in eqs. (13)

$x_t(j)$ = the acreage of stock aged j held to begin time period t grown to age $j+1$ by the end of the period
$h_t(j)$ = the acreage of stock aged j harvested at the end of time t
P_t = the marginal value of harvest volume at the end of time t
$V(j)$ = the merchantable volume per acre from stock age j
C = the per acre regeneration costs
$a_t(j)$ = the marginal flow of amenity service value from stock growing to age j in year t
$\lambda_t(j)$ = the marginal value of stock aged j to begin time t
e^{-it} = the discounting factor for t years, with a continuous interest rate i

unstocked land the necessary condition is:

$$\max\{\left[(a_t(1)+\lambda_{t+1}(0))e^{-i}-\lambda_t(0)\right]; \\ \left[(a_t(1)+\lambda_{t+1}(1))e^{-i}-C-\lambda_t(0)\right]=0, \tag{13a}$$

while for stocked land ($j=1,\dots,t+N-1$):

$$\max\{\left[(P_tV(j+1)+a_t(j+1)+\lambda_{t+1}(0))e^{-i}-\lambda_t(j)\right]; \\ \left[(a_t(j+1)+\lambda_{t+1}(j+1))e^{-i}-\lambda_t(j)\right]\}=0. \tag{13b}$$

In each line the first bracketed value is the return associated with a marginal increase in harvesting, the second value is the return to allowing a marginal increment of stock to go unharvested (regeneration, if $j=0$). The chosen action is to be that which provides the greatest discounted marginal return among these two choices. The new variables $\lambda_t(j)$ may be interpreted as the marginal value of an increment of land with stock aged j. That is, for all t

$$\lambda_t(j)=\partial J(x_t)/\partial x_t(j). \tag{14}$$

Table 12.1 provides a summary of definitions for the various terms in eqs. (13).

This manner [eq. (13)] of presenting the necessary conditions for optimality is useful for highlighting the equivalence of the welfare-maximizing solution to an idealized competitive market allocation.

4.3.2. *A market analogy*

Suppose a landowner could receive the unit price $a_t(j)$ for the amenity services of each acre of his land growing to age j. Suppose an active land market existed allowing the transfer of land stocked with age class j at price $\lambda_t(j)$ or, at time $t+1$, at price $\lambda_{t+1}(j)$. Consider P_t to be the given market price for stumpage and i to be the market rate of return on alternative investments, with this rate equal to the social rate of discount.

Let us look at the competitive forces in such a landholding market. In each time period current landholders will choose land uses that give them the maximum discounted return. Of course, an owner will not just focus on his immediate flow of revenue but also will consider the changing asset value of his land. Suppose the best such discounted return, as given on the left side of eqs. (13), were less than zero. Owners would then sell lands. They would do better putting the proceeds into the market investments which grow in value at rate i. (The present discounted value of the return to such an investment, net of the initial purchase price, is of course zero when discounting is at the same rate i.) The result of the land sales would be to drive down the current price of land.

On the other hand, suppose that at current prices, owners of land in some age class were to make positive returns over the period, other investors would enter the land market and bid up the price of this land. The market can be in equilibrium only when the current price of land for each age class, in each time period, is such that the net discounted return is exactly zero. That is, the equality in eqs. (13) must hold. In other words, the investment in timberland of any age should return exactly at the market rate in each time period.

Simultaneously with the price adjustment for land, bringing equilibrium in the asset-holding market, we would have adjustment in the product markets. In the final equilibrium, stumpage price P_t must equate supply and demand in each period; the payment for the asset services of the stock would be such that if each individual paid his marginal value share of the total marginal value per acre of stock of age j, no individual would want the age mix of the stock altered, and there would be no excess supply or demand for timberlands. The prices P, a, and λ would sustain the welfare-maximizing solution.

Our maximization problem can be viewed precisely as searching for such a market allocation, correcting for the lack of price signals which would have aided the profit-maximizing producers in choosing an appropriate level of amenity services. Conditions (13) can be viewed as a statement of the idealized competitive land market equilibrium.

5. The harvest solution

In this section we derive some results that are useful for describing the pattern of harvests and stock holdings under an economic solution. We focus largely on the nature of the long-run solution. The long-run solution is contrasted to the forester's ideal of the even flow, maximum sustained yield forest.

Under our assumption that demand and productivity are unchanging over time, we might anticipate that the forest area would eventually settle into a stable pattern of supply, a steady state. One example of such a steady state is the even flow forest in which the forest repeats identically in each period with an equal distribution of age classes and with the oldest stands cut each year. It is shown that the multiple use forest need not converge to an even flow state. Rather, periodic cycles with fluctuating harvest levels may be preferred. Furthermore, if an even flow state is appropriate, it may not resemble that typical of a timber management solution. We may find some areas maintained as clearings and other areas maintained as old growth habitat. We may choose to allocate some fraction of the area to a short even flow rotation and another fraction to a long even flow cycle. If an even flow state does result, the harvests will be at a Hartman age and will not generally provide the maximum average yearly flow of value.

The approach taken in the next section is to use the necessary conditions for optimality given by eqs. (13) to characterize the optimal timing of a harvest. This provides us with a generalized version of the Hartman harvesting condition [see eq. (4)]. In Section 5.2 we investigate the long-run steady states which are consistent with the optimal harvest timing condition. In Section 5.3 the more general characteristics of the harvest solution are described. It proves convenient to consider the time period to be so short that we may consider time continuous. This allows direct comparison to the earlier Hartman results as well as a simpler exposition.

5.1. The harvest timing condition

Suppose stocks on land regenerated at time k were optimally harvested at age T, at time t. The marginal value of unstocked land at time k can be represented as[23]

$$\lambda_k = P_t V(T)\mathrm{e}^{-iT} + \sum_{n=1}^{T} a_{k+n}(n)\mathrm{e}^{-in} - C + \lambda_t \mathrm{e}^{-iT}.$$

This expression gives the value of the marginal increment of land as equal to the present value of the flow of amenity and harvest value from a single harvest cycle, plus the present value of the harvested land held at the end of that cycle. With continuous time, the analogous expression is given by

$$\lambda_k = P_t V(T)\mathrm{e}^{-iT} + \int_0^T \left(a_{k+n}(n)\mathrm{e}^{-in}\right)\mathrm{d}n - C + \lambda_t \mathrm{e}^{-iT}. \qquad (15)$$

Optimality of the harvest age T requires that we should be indifferent to a marginal delay in the harvest date. That is, the derivative with respect to t of the right-hand side in eq. (15) should equal zero (remember harvest age is a function

---

[23] The given condition, perhaps obvious, can be found by combining successive necessary conditions for the period of growth and for the final harvest date. To simplify notation we have used λ_t to represent $\lambda_t(0)$, the marginal value of unstocked land at time t.

of date t with $T = t - k$). We find [24]

$$\dot{P}_t V(T) + \dot{\lambda}_t + \left[P_t V'(T) + a_t(T) - i(P_t V(T) + \lambda_t)\right] = 0. \tag{16}$$

Equation (16) is a rather more general version of the Hartman harvesting condition [compare eq. (4)]. Simply, if one holds timber uncut, one had better do at least as well as if one sold the land and stumpage, investing the proceeds at the market rate of return i. The benefits of delaying a harvest, in addition to the value from incremental timber growth and the flow of amenity value over the period of delay, include the possible benefits from a change in stumpage price or from a change in the value of the land itself. Using a condition much like eq. (16), Lyon (1981) discusses the time path of timber prices. Multiple use values were not an issue in his discussion.

5.2. *The steady state solution*

Let us define a forest area as being in a steady state if it is in a repeating cycle with identical stock conditions and harvest levels occurring each T years. What we wish to do is to characterize the steady states that are possible long-run solutions to the multiple use management problem. Can there be fluctuating harvest levels? How does the harvest age compare to the Hartman multiple use rotation age?

Suppose the forest is in a steady state with some cycle T. For this to be a potential solution to the forest management problem the optimality condition (16) must be met. That condition characterizes optimal rotation ages and the corresponding time path of prices at all times, including an eventual steady state. Suppose, however, that condition (16) was not consistent with prices repeating cyclically with the same period T. Then our assumption that T was a possible steady state solution must have been incorrect. We are looking for those values of T for which eq. (16) is consistent with prices following the assumed steady state pattern, repeating cyclically each T years. In some particular cases we will find that the only possible cycle will further require that the forest is unchanging over time, an even flow condition.

In the steady state, stumpage price, the shadow prices of amenity services of stocks, and the land value must repeat cyclically with period T so that $P_t = P_{t+T}$, $a_t(n) = a_{t+T}(n)$, and $\lambda_t = \lambda_{t+T}$. If we are in a steady state with harvest cycle T,

---

[24] Derivatives with respect to time are represented by a dot over the variable. So, $\dot{P}_t$ represents $\mathrm{d}P_t/\mathrm{d}t$.

the marginal value of land held at time t can be expressed as

$$\lambda_t = \left[P_t V(T) \mathrm{e}^{-iT} + \int_0^T \left(a_{t+n}(n) \mathrm{e}^{-in} \right) \mathrm{d}n - C \right] \Big/ (1 - \mathrm{e}^{-iT}).$$

The change in this value over time is given by

$$\dot{\lambda}_t = \left[\dot{P}_t V(T) \mathrm{e}^{-iT} + \int_0^T \left(\dot{a}_{t+n}(n) \mathrm{e}^{-in} \right) \mathrm{d}n \right] \Big/ (1 - \mathrm{e}^{-iT}).$$

Substituting these two equations into the general optimality condition (16) gives us a necessary condition which must be met by any optimal steady state solution. In particular, we find that the steady state must satisfy:

$$\begin{aligned} &\left[\dot{P}_t V(T) + \int_0^T \left(\dot{a}_{t+n}(n) \mathrm{e}^{-in} \right) \mathrm{d}n \right] \\ &\quad = i \left[P_t V(T) + \int_0^T \left(a_{t+n}(n) \mathrm{e}^{-in} \right) \mathrm{d}n - C \right] \\ &\quad - \left[P_t V'(T) + a_t(T) \right] (1 - \mathrm{e}^{-iT}). \end{aligned} \tag{17}$$

This equation relates changes in prices over time to current prices and to the rotation cycle T. The price changes implied by this equation must be consistent with the assumptions of the steady state. That is, the prices must repeat with cycle T.

Before proceeding, we note the correspondence between the right-hand side of eq. (17) and the Hartman harvesting condition (4). It is apparent that this expression, on the right, will have zero value if the harvest cycle T equals the Hartman age associated with the prices P_t and $a_{t+n}(n)$. The value will be less than zero if the harvest cycle T is shorter than the Hartman age and greater than zero for a cycle T longer than the Hartman age.[25] Using these results, we characterize the steady states which are consistent with optimality.

5.2.1. Even flow steady states

Quite clearly there is at least one possible steady state, an even flow steady state at a Hartman age. Furthermore, there can be no even flow steady state except with harvests at a Hartman age. This is consistent with some area being set aside as either old growth or as clearings.

Let us see why a Hartman age is required for an even flow steady state. Suppose the harvest cycle is not a Hartman age. Then the right-hand side of eq. (17) is not equal to zero. This would imply that at least some price must be

---

[25] Second-order conditions to the Hartman problem give this result. Some complication is introduced if there are multiple solutions to the Hartman problem but the basic conclusion is unchanged.

changing at time t. But, that would be inconsistent with the assumption of an even flow steady state. With an even flow condition all prices must be constant over time. Only if harvests are at a Hartman age could all prices be constant. What is perhaps not immediately apparent is that there are many potential even flow steady states. Typically one thinks of an even flow as requiring the whole forest be managed for timber. In fact, what we may find is that the steady state solution calls for the set-aside of some fraction of the forest in either clearings or old growth. We may also find fractions of the managed forest area are harvested at different rotation ages. This is possible because there may be several Hartman ages among which we are indifferent. Such set-asides are still compatible with the forest repeating identically and the harvest flow being constant.[26]

As the proportion of the forest in old growth or clearings decreases we might expect their marginal value to increase, perhaps sufficiently so that some preservation is desirable. Long-run solution with set-asides of clearings or old growth will occur if these age classes have particularly high values at acreage levels not easily met by a fully managed timber forest. For example, setting aside a clearing and managing the remaining lands on a timber rotation might be superior to managing the full forest on a very short rotation to provide the same percentage in clearings but little or no commercial timber. Similarly, setting aside old growth while managing the remainder of the area on a short timber rotation will usually be preferable to choosing extremely long rotation ages which tie up large amounts of stock at low rates of return to achieve the same proportion of old growth.

5.2.2. Cyclical steady states

In the general multiple use problem there is a possibility of a steady state with cyclical fluctuation in the harvest level. This is in contrast to the case in which timber alone is valued. For that special case it is clear that the only possible steady state is an even flow forest harvested at the Faustmann rotation age. It is helpful to describe this timber case in some detail.

When timber values alone are considered eq. (17) becomes:

$$\dot{P}_t V(T) = i(P_t V(T) - C) - P_t V'(T)(1 - \mathrm{e}^{-iT}). \tag{18}$$

This is a simple first-order linear differential equation which may be explicitly solved for the time path of prices. The right-hand side of eq. (18) should be compared to the Faustmann condition (2). It is apparent that only with T equal to a Faustmann age is it possible for prices to repeat periodically. In fact, price must be constant. For any other rotation cycle, the price path implied by eq. (18) is either steadily increasing or declining. Let us see why this is so.

---

[26] We must suppose the amenity services of old trees approach a constant level independent of further increases in age.

Suppose T were greater than the Faustmann age appropriate for a price at the level P_t. In that case, the right-hand side of eq. (18) is positive (see Figure 12.1) and the current price must be rising. In fact the price must continue to rise. To see this, we need to note that the Faustmann age associated with the subsequent higher price level is certainly lower than T. Remember that higher price levels lead to shorter Faustmann rotation ages. Therefore the righ-hand side of eq. (18) remains positive in the next period and price must continue to increase. This steadily rising price is inconsistent with the steady state assumption of cyclically repeating harvests. Similarly, for T less than the Faustmann age associated with price level P_t we can show that eq. (18) implies steadily declining prices and cannot be consistent with the harvests repeating in a steady state cycle. With T equal to the Faustmann age, the right-hand side of eq. (18) equals zero. The timber price is constant at all times and thus, not only does the forest repeat periodically, there must be an even flow of harvest. Only the even flow steady state at the Faustmann age is possible.[27]

With timber alone valued we will have steady states with set-asides of old growth or unregenerated clearings only if timber management is unprofitable. In the extreme case, we may leave inherited stocks unharvested because harvest costs exceed revenues. Once cut, lands would be left unregenerated and unmanaged in the steady state only if the management costs equaled or exceeded the present value of the net harvest revenue.

In the multiple use case we cannot be sure that an even flow steady state is the only possibility. Cyclical steady states with fluctuating harvest levels are possible, with the period of the harvest cycle, in some sense, an average Hartman age. Let us follow the same argument used for timber and see why the general case may differ.

We have the differential eq. (17) which describes a composite change in stumpage price and nontimber shadow prices as depending on the rotation age T. The right-hand side of eq. (17) will equal zero if the steady state cycle T is equal to the Hartman age associated with current prices P_t and $a_{t+n}(n)$. For T greater (less) than this current Hartman age the right-hand side of eq. (17) is positive (negative).

Let us suppose the steady state harvest cycle T is initially greater than the current Hartman age. We know then that the composite set of prices must be rising (although not all the individual prices need rise). In contrast to the timber case we cannot be sure that this implies the composite price level will continue to rise in subsequent periods. In particular, if the price rise is concentrated in increased shadow prices for older stock (perhaps coupled with falling marginal values for younger stocks), then we may find that the Hartman age associated

---

[27] This conclusion does depend on our assumption of continuous time. With discrete time, it can be shown that the long-run solution may have cyclical fluctuation in harvest levels. The shorter the time interval and, generally, the lower the interest rate the less the deviation from the even flow cycle that is possible. See Mitra and Wan (1981a, 1981b) for further discussion of the timber harvesting solution.

with the subsequent price levels has gradually risen until it is above T.[28] At this point the general price level would begin to decline. Now, if the decreases in the shadow values of older stock gradually becomes the dominant change, the rotation age might again fall below R and the price level rise again. Clearly a cyclical pattern of prices consistent with a steady state is a possibility.

It is not too surprising that cyclical, uneven flow management may be desirable. Simply it may be better to receive a high value periodically than little value steadily. The examples that come to mind are related to wildlife. In the early part of this century forest fires in Idaho and Montana opened up large areas of the forests to young vegetation and browse. The result was a tremendous expansion in the elk population and these forests became prime hunting areas for elk. When such extreme conditions cause a sufficiently large gain in value compared to any more moderate even flow condition then periodic heavy harvesting may well be a preferred alternative. Clark (1976) has some discussion on "pulse" fishing. These economics of scale in costs were associated with periodic heavy fishing. Here a range of increasing marginal benefits, for example a threshold level of clearings before hunting is worthwhile, would seem to be necessary for such periodic management.

5.2.3. *Multiple steady states*

We should note that, once the full range of multiple use values is considered, there may be many stable steady states consistent with a supply and demand equilibrium. The initial condition of the forest will determine which of these is approached. This will be particularly noticed with the old growth set-asides. If a forest is inherited with relatively young growth it is most unlikely that we would forego the harvest of commercial timber for the perhaps 200 or more years needed to develop an old growth habitat. The current value of the timber is probably high and the potential benefits of the old growth habitat far in the future. However, if the same land were inherited with old growth stock we might well find it best to save some of the highly valued old growth habitat.

5.2.4. *The steady state supply*

In this section we consider how the long-run supply of timber may be influenced by multiple use management. The greatest average flow of timber is produced at the harvest age $\hat{T}$ which solves the problem[29]

$$\max_{T} \{ V(T)/T \}. \tag{19}$$

---

[28] See the discussion in part 3 on comparative statics of the Hartman solution.

[29] The earlier discussion of maximum yield in part 3 was in terms of average net yield $(PV(T) - C)/T$.

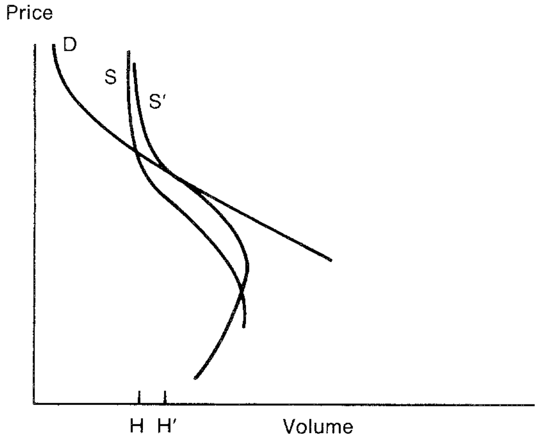

Figure 12.2. Long-run harvest supply.

The solution must meet the first-order condition $V'(T) = V(T)/T$ with harvests at the age at which the increment in growth equals the average harvest flow. Under management for timber alone, the long-run harvest rotation will almost certainly be shorter than the age at which yield is maximized.[30] The steady state economic timber supply will then be less than the maximum attainable. When amenity values are considered the steady state harvest age may well be greater than the Faustmann timber age and the average timber yield greater than under management for timber alone. It is an empirical question as to whether this would lead to a long-run rotation age close to the age of maximum yield.

The effect of introducing multiple use values into the economic problem may seem a little surprising. Suppose that the amenity services were favored by older stands so that the (average) Hartman rotation age in the steady state exceeded the corresponding Faustmann age. As long as both these rotation ages are below the age of maximum physical yield, the multiple use steady state will provide a higher average flow of timber. That is, there will be a higher average flow per acre from those lands under harvest management. We may also find the multiple use values justifying management on more lands, further increasing the supply. However, any set-aside of clearings or old growth excluded from timber management could obviously reduce supply.

In Figure 12.2 we illustrate hypothetical long-run curves for timber from land of uniform quality. The curve labeled S corresponds to the case in which timber

---

[30] The longest possible Faustmann rotation is at T such that $V'(T)/V(T) = i$. This solution occurs when price is so low relative to costs that the maximized land value equals zero. Only if $i < 1/\hat{T}$ is there a (low) timber price that would give a Faustmann age greater than the maximum yield age $\hat{T}$.

value alone is considered, the curve S' represents timber supply when amenity values have been considered. We may find that timber is supplied at lower price levels under multiple use management – the combined benefits justifying management. At high timber prices, as the timber value becomes sufficiently dominant, there is little difference in the timber supplied under the two cases.

The shape of the supply curves may be explained as follows. The slope of the supply curve is given by the derivative

$$\frac{\mathrm{d}(V(T)/T)}{\mathrm{d}T}\frac{\mathrm{d}T}{\mathrm{d}P}=\frac{V'(T)-V(T)/T}{T}\frac{\mathrm{d}T}{\mathrm{d}P}.$$

The slope thus depends on the sign of $\mathrm{d}T/\mathrm{d}P$, the change in the rotation age with respect to an increased steady state price level, and on the sign of the term $(V'(T)-V(T)/T)$, which we have seen equals zero at the age of maximum yield. For ages younger than the age of maximum yield the value $V'(T)-V(T)/T$ is greater (less) than zero. That is, the average harvest is increasing up to the age of maximum yield and declining for longer rotations. The sign of $(\mathrm{d}T/\mathrm{d}P)$ has been discussed in the earlier section on the Faustmann and Hartman models.

When timber alone is considered we know that a higher price level would shorten the rotation age $(\mathrm{d}T/\mathrm{d}P<0)$. For sufficiently low timber price levels relative to costs we might have long economic timber rotations perhaps exceeding the maximum yield age. In this price range the long-run supply of stumpage would be increasing with higher price levels. Once the price is sufficiently high so that the rotation is below the maximum yield age, the long-run supply begins to decline with the shorter timber rotations that result from higher price levels. At high prices, the supply approaches a constant level as the rotation age approaches age T_0, the Faustmann age when costs are zero.

As long as we may suppose that the multiple use rotation age, at any given price, is longer than the timber rotation, then, again, higher timber prices lead to shorter rotations. As a result, the shape of the timber supply curve under multiple use management is much the same as under timber management alone. The relative position of the two supply curves in Figure 12.2 is explained by the longer multiple use rotations. For example, at higher price levels, with rotations shorter than the maximum yield age, the multiple use rotation leads to a greater average harvest than is supplied under the shorter timber management rotations. In this case, there is a range of prices for which the timber supply under multiple use management exceeds the supply under management for timber alone.

Superimposed on the long-run supply curves in Figure 12.2 we have drawn a demand curve labeled D. With such a demand curve, constant over time, the long-run timber harvests would be the amounts H or H' for the timber and multiple use case respectively, where the demand crosses supply curve. In both cases these harvest levels, as drawn, are less than the maximum sustainable yield. Harvesting at the age of maximum yield will only coincidentally correspond to the long-run economic policy. Even then such a policy would provide little guidance for harvesting in earlier years.

5.3. *The approach to the steady state*

Equation (16), along with the conditions of market clearing, can be used to describe a time path of prices, harvests, and harvest age over time. To do so explicitly is rather formidable, even for the case when timber alone has value.[31] There is a strong indication that with constant demands there is a gradual convergence to a steady state such as we have described earlier. Prices will generally be rising in periods in which stands older than the current Hartman age are being cut and prices will generally be falling in periods in which we must harvest younger growth. The relative rate of price change over time will not generally exceed the discount rate, although it may if inherited stocks of old growth are being cleared. The price fluctuation will gradually dampen. The harvest ages in the earlier periods may differ greatly from the steady state Hartman age.

It is worth giving a few basic conclusions related to the harvest during the transition period to a steady state. In doing so we can highlight the difference in the nature of the solution from solutions to models in which the services of each stock class are treated as independent of the mix of stocks. The optimization of the general multiple use model leads to a set of shadow prices for stumpage, for the services of each age class of standing stock, and for land in each age class. Suppose we were given these prices, we could use a simple recursive procedure to solve eqs. (13) for the optimal solution. The shadow prices decentralize the harvesting decision, allowing the optimal policy for each stand to be evaluated independently of its neighbor. In this sense, linear programming approaches used by the U.S. Forest Service may be justified. The difficulty is that the shadow prices cannot easily be determined without having explicitly solved the more general problem. There is no market which reveals value and these prices themselves are determined by the full mix of standing stock.

The shadow prices of standing stock depend on the distribution of age class in each time period and change as this distribution is altered. As a result of this shifting pattern of prices, the harvest timing rules can be very complex, not easily reflected by any rule of thumb. Even when the price of timber is constant, it is unlikely for a stand to be harvested at the same age for two successive rotations. Only when the identical forest condition exists is the same harvest decision made. The harvest decision in each time period balances the improvement in current value against the future benefits arising from adjusting the distribution of age

---

[31] See Heaps and Neher (1978) and Lyon (1981) for related discussion of the timber case.

classes. At any particular time, we may find the solution calling for delaying regeneration on some areas, forgoing harvests on some older stands, and harvesting other younger stands.

Because of the complex nature of the multiple use harvest solution it is helpful to provide a few examples which illustrate some of the more unusual characteristics of the solutions. In Section 6 these examples are described.

6. Policy considerations and illustrations

In the earlier sections of this chapter we have considered, in a rather abstract style, a number of the controversial issues related to multiple use management of public forestlands. In this section we will provide some illustrations to highlight the policy concerns. The examples use fairly realistic data to illustrate the range of solutions we may find under an economic approach to multiple use management. In Section 6.1 the formulation of the examples is described. In Section 6.2 a set of examples is described. The examples are selected to illustrate the effect of multiple use values on the decision to manage an area, the age of harvest, the diversity of the forest habitat, and the level of timber supply.

6.1. *Formulation of examples*

We consider the harvest decisions on a set of interrelated subareas of a forest unit. The value of the amenity services from the full unit depends on the mix of ages across these subareas. The examples are solved as simple dynamic programming problems. The basic recursion is essentially that given by problem (12). To reduce the dimensions of the problem it is assumed that there are a small number of stands, that management decisions are limited to either harvesting or growing the full acreage in each stand, and that there is a fairly large time interval between harvest decisions. Specifically, there are eight stands of equal area and productivity in the forest unit and we consider age in 20-year increments, up to age 100. If trees are left uncut at age 100, we assume that there is no further growth and no mortality. One may think of the "stands" as being dispersed across the forest unit in patches of appropriate size. Nevertheless, our harvest decision requires that a whole stand be harvested, rather than allowing any smaller fraction to be cut.

6.1.1. *Timber yields and costs*

The timber yields are based on Douglas fir growth tables for the Pacific Northwest from McCardle and coauthors (1961). Yields are given in Table 12.2. This

Table 12.2
Yields from Douglas Fir stands – one acre, site index 80

Stand age (yrs.)	Managed yields[a] (cu. ft.)
40	2110
50	2840
60	3500
70	4090
80	4580
90	5000
100	5350

[a] *Source*: McCardle and coauthors (1961).

table gives total cubic foot volume per acre as a function of stand age. This is for a stand managed so as to ensure full stocking and no delay in regeneration.[32]

The net price (P) for stumpage varies in the examples over the values \$0.25, \$0.50, and \$1.15 per cubic foot. The price does not vary with the harvest level. The management costs for the fully stocked stands were taken to be \$115 per acre, incurred at the time of planting.

6.1.2. *Amenity values*

Few data on either forest use or forest conditions exist that would allow us to estimate actual demand functions for amenity services. We have chosen to base our values on indices of habitat suitability for wildlife and aesthetic quality described in Boyce (1977, 1978). These indices depend on the mix of stand ages and the size of individual stands. Essentially, three factors determine our amenity value. It is most desirable to have some small percentage of the area clear, a larger percentage of the area in old growth, and a fairly balanced age mix with somewhat more area in older age classes than in the younger age classes.

In Table 12.3, column 1 lists a selection of age mixes. For example, the age mix 11112222 represents an area with four stands in age class 1 (growing to age 20) and four stands in age class 2 (age 40). Column 2 gives the corresponding yearly flow of recreation value per acre. For convenience, we assume the ordering of the age mix does not alter the value. It can be seen that the incremental value from a stand of a given age is very dependent upon the age mix of the remaining areas.

---

[32] With approximately 55 cubic feet per acre per year growth, such land qualifies as site productivity class IV, lower quality commercial timberland. Prime timber lands may yield over 200 cubic feet average growth per year.

Table 12.3
Yearly recreation value per acre

Age mix	Amenity value
11122222	0.6
22222222	5.8
11223344	25.5
11223345	29.2
12234555	65.1
22234555	48.8
12245555	73.4
12345555	81.3
12555555	93.9
55555555	64.5

And, although it is usually preferable to have an older stand over a younger one, this is not uniformly so. The amenity values are perhaps high, but may be appropriate for areas with a potential for significant dispersed recreational use, such as lands adjacent to trails.

6.2. Multiple stand – harvest solutions

In Tables 12.4a and 4b we give time sequences of harvest solutions for a representative forest unit. The age mix of the eight stands is given at 20-year decision intervals. The harvest policy is indicated by underscoring the age class to be harvested. A decision to delay regeneration is indicated when a stand in age class 1 is underscored. We assume that any naturally regenerated stock is not commercially valuable. For comparison, both the multiple use solution and the corresponding timber management solution are given. The harvest policy for the first five decision periods is given.

6.2.1. Timber management solution

The harvest policy under timber management is to cut any initial holdings of stock aged 40 or above. With a discount rate at 4 percent there is no artificial regeneration (planting) if the stumpage price is $0.25 per cubic foot or less. With a discount rate of 7 percent, there is no regeneration with prices of $0.80 per cubic foot, or less. In these cases the land is withdrawn from active management after harvesting the initial endowments of stock. At higher prices, the land is regenerated and each stand harvested at age 40 in a repeating Faustmann cycle.

Table 12.4a
Harvest policies – multiple use vs. timber management ($i = 4$ percent)

			($P = \$0.25$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	55554221	55555331	55551442	55552551
Timber management Age mix	55443111	11111111	11111111	11111111	11111111
			($P = \$0.50$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	15554222	2155333	32155444	43215555
Timber management Age mix	55443111	11111222	22222111	11111222	22222111
			($P = \$1.15$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	11111222	22222333	11111111	22222222
Timber management Age mix	55443111	11111222	22222111	11111222	22222111

Table 12.4b
Harvest policies – multiple use vs. timber management ($i = 7$ percent)

			($P = \$0.25$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	55554221	555555331	55551442	55512553
Timber management Age mix	55443111	11111111	11111111	11111111	11111111
			($P = \$0.50$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	11111111	11111111	11111111	11111111
Timber management Age mix	55443111	11111111	11111111	11111111	11111111
			($P = \$1.15$)		
	Year 0	Year 20	Year 40	Year 60	Year 80
Multiple use management Age mix	55443111	11111222	22222111	11111222	22222111
Timber management Age mix	55443111	11111222	22222111	11111222	22222111

6.2.2. The multiple use solution – the decision to manage

Under timber management we see that at sufficiently low prices all stands are harvested and abandoned with no regeneration. Under multiple use management there is a greater likelihood of regeneration and continuing harvest. At the lowest timber price (\$0.25, $i = 4$ percent) a low level of timber harvest is maintained under multiple use management. A harvest provides an effective means of maintaining clearings for wildlife.

One interesting case illustrates the potentially complex effect of changing relative prices on timber supply and management. In Table 12.4b, at the lowest price we harvest a limited amount of timber largely for the recreational benefits arising from a diverse habitat. At the intermediate timber price there is complete harvesting of initial stock endowments but no regeneration or further management. At this intermediate price the initial endowment is sufficiently valuable to justify complete harvesting. Once harvested, the area is so lacking in diversity (and the timber price sufficiently low) that further management is not justified. At the highest price, the area is perpetually managed under a timber regime.

In our examples the preservation of some areas of old growth is a commonly chosen policy when the timber value is relatively low. Temporary delays in regeneration are also common in the earlier periods. In Table 4a, with the stumpage price at \$0.25, our unit converges to an even flow steady state ($555555\underline{2}1$) with three-quarters of the area in old growth and the remaining area managed on a 40-year harvest rotation. In the long run solution there are not any permanently set aside clearings. In earlier periods, the temporary maintenance of clearings (harvesting stands in age class 1) provides wildlife forage while allowing the development of old growth. If a larger proportion of cleared land had been desirable in our examples, we might then have seen the solution call for permanent clearings. Here the limited harvesting needed to maintain forage areas does not interfere with the maintenance of older growth and diversity.

At the higher timber prices, the multiple use solution corresponds quite closely to the timber management solution. However, even with high timber values we often see some effect on the harvest solution. Particularly in the initial periods, it may prove advantageous to delay a harvest in order to gain some benefits from a better age distribution. In Table 12.4a with the price at \$1.15 we see it is advantageous to delay some harvests until age 60. This imposes little financial cost and slightly improves recreational value.

6.2.3. The age of harvest

It is obvious that there is no easy description of the effect of multiple use values on the harvest age. This is so even for our example, in which older stands are generally preferred. The benefits of maintaining at least one stand clear and some balance in the age mix makes the harvest timing very dependent upon the specific mix of ages in our unit. This is in strong contrast to the simplistic single stand model.

Harvest ages during the initial transition period to a steady state can be very unusual with younger stands cut while older stands are left unharvested. This can be seen in Table 12.4a at the lower timber price. It is also apparent that a higher relative recreational value does not necessarily lead to longer rotations. Compare the long run solutions for prices $0.25 and $0.50 in Table 12.4a. We see that with the higher timber price, the unit converges to an even flow steady state with three stands in old growth and the remainder harvested on a 100-year cycle. At the lower timber price the harvest is at a 40-year cycle. The increased set-aside old growth compensates for the shorter rotation cycle on the remaining area.

6.2.4. Diversity of habitat and specialization of land use

It is apparent from the examples above that recreational or other multiple use values which depend on the diversity of the forest condition can have a great effect on the harvesting decision. The examples make clear the dependence of the harvest age and the decision to manage on the current mix of age classes. With sufficiently high value attached to the nontimber services the improvement of the age mix may completely motivate the timing of the harvest. The complexity of the harvest solution is due to the nonlinearity of the nontimber benefit function. Such nonlinearity seems generally to be expected.

In the longer run we find that the benefits from diversity of standing stock may lead us to allocate areas of the forest to specialized purposes. For example, we may choose to preserve some areas as old growth. One interesting option is the specialization of use over time rather than location. With the highest timber price in Table 12.4a we choose periodic high timber harvests with intermediate periods of no harvest and moderate amenity value. Such a noneven flow policy, likely to be sensible in many areas, is chosen because of the very low amenity value when more than two stands are cleared.

The eventual steady state can be highly dependent upon our initial diversity of age classes. The most striking example is given in Table 12.5 in which unit 2, initially all old growth, is cleared and not regenerated while the other area, more diverse initially, is managed largely for recreation with a moderate sustained flow of timber harvest. There is a strong penalty on the lack of diversity and balance.

6.3. Conclusion

The general multiple use harvesting policy is seen to be complex. No simple rule of thumb is likely to describe the harvest. We see that sometimes younger stands are harvested, leaving older ones uncut. We may choose to briefly delay regenera-

Table 12.5
Harvest policies – dependence on initial conditions

	($P = \$0.25$, $i = 7$ percent) Year 0	Year 20	Year 40	Year 60
Multiple use management				
Unit 1				
Age mix	$5333221\underline{1}$	$\underline{5}4443321$	$1\underline{5}55432$	$21\underline{5}55543$
Unit 2				
Age mix	$\underline{55555555}$	$\underline{11111111}$	$\underline{11111111}$	$\underline{11111111}$

tion. We rarely cut a particular stand at the same age twice in succession during the initial periods. The forest areas may be managed with some areas set aside for specialized purposes – old growth or clearing for wildlife. We may choose to specialize over time with the land producing high timber yields in some periods and high recreation benefits in others. Even flow policies are not inherently desirable long-run goals. The optimal harvest age is unlikely to be at the age of maximum sustained yield. Indeed, long even flow rotations, far from being the desirable compromise policy for multiple use management, may simply provide both uneconomic timber and a poor balance of age classes for nontimber use. Perhaps most importantly we see, from these examples, that the harvesting decision can be extremely sensitive to factors about which we have little empirical knowledge.

References

Alston, R.M. (1972) *Forest-Goals and Decisionmaking in the Forest Service*, Research Paper INT-128 (USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden).

Alston, R.M. (1979) "Economic Tradeoffs of Multiple Use Management", in: D.D. Hook and B.A. Dunn (eds.), *Multiple-Use Management of Forest Resources: Proceedings of the Symposium* (Department of Forestry, Clemson University, Clemson).

Behan, R.W. (1967) "The Succotash Syndrome, or Multiple Use – A Heartfelt Approach to Forestland Management", *National Resources Journal* 7, 473–484.

Berck, P. (1979) "The Economics of Timber: A Renewable Resource in the Long Run", *The Bell Journal of Economics* 10, 447–462.

Berck, P. (1981) "Optimal Management of Renewable Resources with Growing Demand and Stock Externalities", *Journal of Environmental Economics and Management* 8, 105–117.

Boyce, S.G. (1977) *Management of Eastern Hardwood Forests for Multiple Benefits (DYNAST-MB)*, Research Paper SE-168 (USDA Forest Service, Southeast Forest Experiment Station, Asheville).

Boyce, S.G. (1978) *Management of Forests for Timber and Related Benefits (DYNAST-TM)*, Research Paper SE-184 (USDA Forest Service, Southeast Forest Experiment Station, Asheville).

Bradford, D.F. and G.G. Hildebrand (1977) "Observable Preferences for Public Goods", *Journal of Public Economics* 8, 111–131.

Brown, T.C. and D.R. Carder (1977) "Sustained Yield of What?, *Journal of Forestry* 75, 722–723.

Calish, S., R.D. Fight and D.E. Teeguarden (1978) "How Do Nontimber Values Affect Douglas-Fir Rotations?", *Journal of Forestry* 76, 217–221.

Carlton, D.W. (1979) "Valuing Market Benefits and Costs in Related Output and Input Markets", *American Economic Review* 69, 688–696.

Clark, C. (1976) *Mathematical Bioeconomics* (Wiley, New York).

Downs, A. (1976) "Sustained Yield and American Social Goals", in: B. Dowdle (ed.), *The Economics of Sustained Yield Forestry: A Symposium* (University of Washington, Seattle).

Faustmann, Martin (1849) "On the Determination of the Value Which Forest Land and Immature Stands Possess for Forestry", English edition in: M. Gane (ed.) (1968) *Martin Faustmann and the Evolution of Discounted Cash Flow*, Institute Paper 42 (Commonwealth Forestry Institute, Oxford University).

Gaffney, M.M. (1957) "Concepts of Financial Maturity of Timber and Other Assets", Agricultural Economics Information Series no. 62 (North Carolina State College, Raleigh).

Gregory, G.R. (1955) "An Economic Approach to Multiple Use", *Forest Science* 1, 6–13.

Hartman, R. (1976) "The Harvesting Decision When a Standing Forest Has Value", *Economic Inquiry* 14, 52–58.

Heaps, R. and P.A. Neher (1979) "The Economics of Forestry When the Rate of Harvest is Constrained", *Journal of Environmental Economics and Management* 6, 297–319.

Henderson, J.M. and R.E. Quandt (1971) *Microeconomic Theory*, 2nd ed. (McGraw-Hill, New York).

Hirshleifer, J. (1976) "Sustained Yield versus Capital Theory", in: B. Dowdle (ed.), *The Economics of Sustained Yield Forestry: A Symposium* (University of Washington, Seattle).

Hyde, W.F. (1980) *Timber Supply, Land Allocation, and Economic Efficiency* (Resources for the Future, Johns Hopkins University Press, Baltimore).

Jacobsen, S.E. (1979) "On the Equivalence of Input and Output Market Marshallian Surplus Measures", *American Economic Review* 69, 423–428.

Johnson, K.N., D.B. Jones and B.M. Kent (1980) "Forest Planning Model (FORPLAN) – User's Guide and Operations Manual", Draft (USDA Forest Service, Land Use Planning, Systems Applications Unit, Fort Collins).

Johnson, K.N. and H.L. Scheurman (1977) "Techniques for Prescribing Optimal Timber Harvest and Investment Under Different Objectives: Discussion and Synthesis", Monograph 18, a supplement to Forest Science, vol. 23.

Lyon, K.S. (1981) "Mining of the Forest and the Time Path of the Price of Timber", *Journal of Environmental Economics and Management* 8, 330–344.

Lyon, K.S. and R.A. Sedjo (1981) "Application of Optimal Control Theory to Estimate Long-Term Supply of Timber Resources", Unpublished draft, (Resources for the Future, Washington).

Mäler, K.G. (1971) "A Method of Estimating Social Benefits from Pollution Control", *Swedish Journal of Economics* 73, 121–133.

McArdle, R.E., W.H. Meyer and D. Bruce (1961) "The Yield of Douglas-Fir in the Pacific Northwest", USDA Technical Bulletin 201 (USDA Forest Service, Washington).

McDonough, J.M. and D.E. Park (1975) "Nonlinear Optimal Control Approach to Interregional Management of Timber Production and Distribution", in: J. Meadows, B. Bare, K. Ware and C. Row (eds.), Systems Analysis and Forest Resource Management (Society of American Foresters, Washington).

McKenzie, G.W. and I.F. Pearce (1982) "Welfare Measurement – A Synthesis", *American Economic Review* 72, 669–682.

Mitra, T. and H.Y. Wan (1981a) "Some Theoretical Results on the Economics of Forestry", Working Paper 270 (Department of Economics, Cornell University).

Mitra, T. and H.Y. Wan (1981b) "On the Faustmann Solution to the Forest Management Problem", Working Paper 266, (Department of Economics, Cornell University).

Nguyen, D. (1979) "Environmental Services and the Optimum Rotation Problem in Forest Management", *Journal of Environmental Management* 8, 127–136.

Pearse, P.M. (1967) "The Optimum Forest Rotation", *The Forestry Chronicle* 43, 178–195.

Riiters, K., J.D. Brodie and D. Hann (1982) "Dynamic Programming for Optimization of Timber Production and Grazing in Ponderosa Pine", *Forest Science* 28, 517–526.

Samuelson, P.A. (1976) "Economics of Forestry in an Evolving Society", *Economic Inquiry* 14, 466–492.

Thomas, J.W. (1979) "Wildlife Habitats in Managed Forests: The Blue Mountains of Oregon and Washington", Agricultural Handbook no. 533 (USDA Forest Service, Washington).

USDA Forest Service (1978) *The Principal Laws Relating to Forest Service Activities*, Agricultural Handbook no. 453 (USDA Forest Service, Washington).
Walter, G.R. (1977) "Economics of Multiple-Use Forestry", *Journal of Environmental Management* 5, 345–356.
Willig, R.D. (1976) "Consumer's Surplus without Apology", *American Economic Review* 66, 589–597.
Zivnuska, J.A. (1961) "The Multiple Problems of Multiple Use", *Journal of Forestry* 59, 555–560.

Chapter 13

LAND RESOURCES AND LAND MARKETS

ALAN RANDALL and EMERY N. CASTLE*

University of Kentucky and Resources for the Future, Washington

1. Introduction

The economics of land is no simple study. Depending on the human purposes one has in mind, land can usefully be viewed as soil, a store of minerals, terrain, territory, property, or capital asset. Thus, land may serve as an input in agricultural or forest production, wildlife habitats and the support of complex ecosystems, park lands, and suburban and urban landscapes. It may be viewed as a store of minerals and an obstacle to their discovery and recovery. Land may be viewed as terrain, or terra firma, directing the hydrological, atmospheric, and micro climatic systems, and supporting human-made structures such as buildings and transportation and communication systems. Land may be viewed as a capital asset, a marketable store of wealth and an investment opportunity.

Furthermore, one may conceive of land as a fundamental organizing principle for human society: as an arbiter of spacial relationships (the notions of distance, space and territory); a fundamental basis of legal rights and privileges (consider the special place of landed property in public and private law); an essential element in the structure of social relationships (consider for example the social status conferred by land ownership in a wide variety of societies); and an essential component of a minimal set of privileges for the common folk (consider the role of the common lands in many societies, and the widespread practice of regulating land uses in order to provide an acceptable level of environmental amenities for all).

It is little wonder, then, that the economics of land is a challenging, complicated, and sometimes confused subject, and that economists have no monopoly in studying the role of land in human relationships. Land is a central concept, not only in the technical sciences and professions such as engineering, geology, agriculture and forestry, but also in law and every one of the social sciences.

---

*We gratefully acknowledge the detailed research assistance of Olvar Bergland and the helpful suggestions of Bergland and John P. Hoehn.

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

This chapter is mostly concerned with the economics of land. Of course, one can study the economics of literally any commodity or factor of production. In most cases, that is where it ends. The principles of economics are brought to bear on the topic at hand, and that subject-matter is thus illuminated. But the flow is essentially one-way: the content of the theory and principles of economics is itself little changed for the experience.

On the other hand, there exists an ancient and persistent suspicion that land is not just any commodity or factor of production. Thus, land itself played a prominent role in the development of general economic theory and retains a special position in some current renditions thereof. There is recurrent debate as to whether any economic theory that treats land as "nothing special" can be valid. For this reason, we must be concerned not only with the economics of land but also with the place of land in economics.

1.1. Land in mainstream economic theory

In the most general versions of mainstream economic theory, land is treated as a factor of production and the debate revolves around whether it is useful to reserve a special place for land in formulating the aggregate production function.

The classical economists were writing in Western Europe (and primarily England) as the first phase of the agrarian and industrial revolutions was drawing to a close. Their experience of everyday life made them accustomed to technological progress and aware of the importance of capital to productivity. Nevertheless, in this era of the enclosures and corn laws, industrial progress had as yet done little to improve the lot of the increasingly urbanized common people, and the threat of famine was ever-present.

The classical economists conceptualized the aggregate production function as taking the form:

$$Y=f(D,K,L),$$

where $Y=$ aggregate output, $D=$ land, $K=$ capital, and $L=$ labor. Land was broadly defined as synonymous with "the natural endowment", and could thus be interpreted as meaning the totality of natural resources, as natural resources were understood at the time. It was thought essential to consider land separately from the other factors of production. Land was treated as fundamentally fixed in availability, while the supplies of capital and labor were surely more elastic. Land as the fixed factor provided the rationale for diminishing returns from other inputs,[1] and gave rise to the notion of Ricardian rent, one of the more durable

---

[1] The concept of diminishing returns is clearly a precursor to the marginalist economics of one-half century later. Interestingly, it was Thunen, celebrated for his theory of location and land use, who enunciated in fully developed form the principles of marginalism at a time when the better known economists of his day were unable to do so [Samuelson (1983)].

among the numerous concepts of economic rent that have been advanced in the literature. As population growth outstripped improvements in agricultural technology, and progressively poorer land was pressed into agricultural production, economic surpluses or rents would accrue to land of higher quality.

These same notions – the crucial role of land in producing food and raw materials, the fixity of land and consequent diminishing marginal productivity of other inputs in agricultural production, the limited responsiveness of food production to new technologies, and the consequent increasing relative scarcity of land – underlay Malthus' pessimistic prediction that human populations would always be restrained by the carrying capacity of the agricultural system, thus dooming the great mass of human society to a mere subsistence standard of living.

The milieu of the neoclassical era (which started about a century after the classical era began) was quite different. The industrial revolution was more than one hundred years old, with no end in sight. This longevity, together with evidence that the pace of technological improvements in productivity was if anything increasing, gave rise to increasing confidence in the power of technology to improve the lot of humankind. Three factors – continued technological developments in agriculture, substantial immigration from Europe to the New World, and the increasingly large and reliable flows of food and raw materials from the rest of the world to Europe – had provided some considerable respite from the Malthusian specter, at least in the industrialized countries. These economies were no longer dominated by food and fiber industries, but had diversified toward mineral materials and fossil fuels. It seemed that the key to material progress lay in the accumulation of capital and, what is more, each new generation of capital equipment embodied a new and superior technology.

To neoclassical economists, the fixity of land and diminishing marginal productivity of the other factors were no longer the dominant economic reality. For many purposes, neoclassical economists tended to give land no special place in their models. The neoclassical aggregate production function could reasonably be said to read:

$$Y = g(K, L).$$

The point is not that land no longer seemed to matter. Rather, there seemed no reason to accord land any special treatment that would suggest its role is quite distinct from that of the other factors.[2] Land could safely be subsumed under the

---

[2] The neoclassical growth models of Solow (1956) explicitly use the $Y = g(K, L)$ production function. His well-known article (1974) enquiring about the long-term prospects for an economy that uses exhaustible natural resources posits a production function of the form $Y = f(D, K, L)$. However, the treatment of natural resources as deserving special consideration is more apparent than real. Solow goes on to use a production function having the property of constant unitary elasticity of substitution among inputs, which obviously undermines any claim of natural resources to uniqueness in the production process.

broader aggregate of capital, since (i) its productivity was clearly responsive to investment and the application of technology, and (ii) the increasing economic importance on non-food-and-fiber commodities together with the increasing use of capital inputs in even the food and fiber industries suggested very substantial possibilities for substitution between land and capital.

For the current generation of economists, everyday life experiences have left vastly different impressions from those that formed the attitudes of the classicals. For the ordinary citizens of post-industrial countries, prosperity is the norm, technological progress pervades the whole economy from farms to the information industries, and the acquisition of education and training is a lifelong process.

A more recent perspective, initially associated with the Chicago School, recognizes no fundamental distinction between labor and capital in the aggregate production process. Labor productivity is seen as responsive to investment in training and education, and through this process labor embodies new technologies just as do other forms of capital. From the "human capital" perspective,[3] the aggregate production function could well read simply:

$$Y = h(K).$$

In this formulation, K is given a very modern interpretation. Capital is seen as whatever is created or improved by the act of investment and thus includes physical plant, educated human minds and bodies, farms and forests that respond to investment and management, and the technologies embodied in all of these productive facilities. Furthermore, these various forms of capital are seen as good substitutes, one for another. The point is that investment is considered the only fundamental limitation to the capacity of the human population to support itself on this earth. Natural resource limitations, according to this viewpoint, are simply not fundamental.[4] They can be overcome by substituting capital, physical and human, for limited natural resources.

We do not mean to leave the impression that all modern economists take the cornucopian perspective. During the 1970s, the environmental crisis, the oil embargo and the emergence of OPEC as an effective price-setter encouraged perhaps a minority of modern economists to take a neo-Malthusian perspective. The basic laws of physics were seen as limiting economic opportunities for humankind. Nevertheless, the mainstream of economic thinking, as it has developed in the more prosperous countries over the last two centuries, has tended to

---

[3] The seminal statements of the human position are attributable to T.W. Schultz. A currently popular exposition of the argument that human capital is essentially limitless in its capacity to substitute for natural resources is attributable to Simon (1981).

[4] Consider Schultz's Nobel Lecture (1980) in which he reiterates the case that land is overrated and Ricardian rent has lost its sting.

de-emphasize the fixity of land and natural resources while performing broad and general analysis of aggregate production and economic growth.[5]

While modern renditions of mainstream economic theory are less likely than their classical precursors to reserve a special place for land and natural resources, there is no denying the important role of land in the development of economics itself. The pivotal concepts of diminishing returns to the fixed factor and its consequent ability to accrue rents, and the later development of a full-fledged marginalist microeconomics, surely emerged from economists' perceptions of the significance of land in the production processes of their time. Furthermore, Thunen, the seminal land economist, not only grasped the principles of marginalism that eluded his contemporaries (see footnote 1), but outlined the basis for general equilibrium economics [Samuelson (1983)].

Thus far, we have focused exclusively on land as a factor of production, usually for raw materials. However, the modern reality is that land is equally as important as a consumption good. Economists have not ignored this development, as the modern elaboration of location theory, urban economics and environmental economics attests. And some important contributions to economic theory and method have been made by specialists in these fields. Nevertheless, it is difficult to make a case that these fields of inquiry have yet contributed as much to general economics as did the earlier attention paid to land as a factor of production.

Land and natural resource concepts, of course, remain important for many special purposes in economics. The Ricardian concept of economic rent remains durable, and finds application in areas as diverse as land economics, location theory, and welfare change measurement. Agricultural economics, natural resource economics, urban economics and regional economics are important areas of specialization within economics, and the classical concept of land, appropriately updated, plays an important role in each.

1.2. Land economics

Modern land economics is the product of diverse influences, and some perspective can be gained by briefly tracing these.

1.2.1. The classical economics of land

The economics of Ricardo and Malthus was very much an economics of land. As Samuelson (1959b) astutely observes, the logical import of the Ricardian system

---

[5] Castle (1982) provides a review of the recent and current debate between cornucopians and catastrophists.

leads not to a labor theory of value, as Ricardo himself concluded, but to a land theory of value. While it is possible, with hindsight, to criticize Ricardo's failure to perceive completely the consequences of his concept of rent, that concept itself must be recognized as a major development.

1.2.2. "Land economics 1" and location theory

Ricardo's analysis of the location of agriculture, based on soil quality (or, to simplify a little, fertility), yielded fewer interesting results about location than did Thunen's.[6] The latter's organizing principle was that, if some essential activity such as exchange takes place at a given point, increasing distance from that point imposes increasing costs. Applying this principle to agricultural production, Thunen was able to deduce that different commodities would be produced in distinct zones, and he identified the principle by which commodities were assigned to particular zones. In addition, his simple model yielded clear directional predictions about the relationships of distance from the center and output per acre, labor employment per acre, and land rents.

Thunen's analysis of the location of agricultural production led him to make major contributions, ahead of his time, to economic theory (noted above). Furthermore, his position as the father of regional economics is assured, while his claims to paternity of location theory are enhanced by the fact that modern geography, too, claims him as the founder of that discipline.

Utilizing recent advances in consumer theory – duality, indirect utility, and the expenditure function – modern urban economists have extended Thunen's theory to the organization of the city and, in so doing, to the concept of land as a consumption good.

In their purest (and sparest) forms, the agricultural location models of Thunen and their intellectual descendants in urban location theory generate an interesting set of predictions but provide somewhat unsatisfying descriptions of reality. Topography is reduced to a featureless plain, while people (be they Thunen's farmers or the worker-consumers of urban economics) are homogenized by the assumption of identical utility functions. It is possible, of course, to introduce various elements of reality. However, this process has a rather *ad hoc* flavor and usually entails some sacrifice of generality, but produces little in the way of unexpected theoretical predictions. It has, nevertheless, provided a basis for a variety of interesting empirical applications in regional economics.

1.2.3. "Land economics 2" and natural resource economics

From the late nineteenth century through the mid-twentieth century, there developed and, for a time, flourished a distinct group known as land economists.

---

[6] For English translations of Thunen (1930), see Dempsey (1960) and Hall (1966).

Their attitudes toward land tended to reflect contemporary social conditions in Europe and the alternatives offered by the New World, while their economics owed less to the neoclassical school than to its critics of the historical and institutionalist persuasions.

In Europe, the peasantry was being expelled from agriculture to join either the urban proletariat or the stream of immigrants to the New World. Among the major attractions of the New World were abundant land, social philosophies that saw widespread land ownership as the fount of human dignity and sociopolitical stability, and land settlement policies to match. For the land economists, the abundance of land did not serve to make land less interesting, as it had for the neoclassicals; rather, it gave the New World special status as a place where the ordinary folk might, perhaps for the first time, seek economic independence and personal fulfillment with some realistic chance of success.

While the land economists retained the classical belief in the uniqueness of the land resource, land economics diverged from classical and neoclassical economics in other ways. The importance of land, and hence the subject matter of inquiry, was seen as extending beyond mere economics into law, sociology, and political science. The basic paradigm was evolutionary, and methods of analysis were holistic, historical and empirical. The standard philosophy of land economics gravitated toward an instrumentalism that saw the scholar as not merely an observer of the unfolding social panorama, but an active participant and reformer. In all of these respects, the contrast with the standard neoclassical perspective is obvious.[7]

During the last three decades, an identifiable sub-discipline of natural resource economics has become established. It has, to be sure, been influenced by the land economics perspective. Nevertheless, it is the product of a diverse set of influences, many of them owing more to the neoclassical heritage than to institutionalism. These include mainstream microeconomics, welfare economics and public finance; the more recent property rights and public choice tradition; and intertemporal resource allocation theory, the roots of which go back at least as far as Faustmann (1849) (see also Chapter 12 of this Handbook). Work in natural resource economics has drawn attention to, and contributed toward the resolution of, the special problems of water, the public lands, minerals, energy, biological resources (which are renewable but destructable), and environmental quality.

On the one hand, the classical concept of land as encompassing all of nature that is of economic significance was considered to be entirely too broad a concept for useful analysis. On the other hand, the classical concept of land as a factor of production was considered unduly restrictive. Just as the earlier land economists had recognized the sociopolitical and even ideological significance of land, modern natural resource economists recognize that significant demands for land

---

[7] For a well-known textbook in the "land economics 2" tradition, see Ely and Wehrwein (1964).

and natural resources are direct demands that arise from the desire to enjoy environmental amenities.

1.2.4. Finance and real estate markets

Land, being long-lived, readily marketable, and capable of producing a time stream of goods and services, is not only a natural resource but a paradigm case of a capital asset. Not unexpectedly, the asset pricing model that emerges from finance and real estate theory bears an immediate familial relationship with the net present value model of intertemporal natural resource economics. Nevertheless, finance and capital market theory makes its special contribution to understanding the economics of land, by clarifying the links between real estate and other asset markets, interest and inflation rates and, ultimately, macroeconomic conditions.

1.3. The focus of this chapter: Land markets

With this prologue, it is now essential to choose a more specific theme for the remainder of this chapter. We will focus on land markets, first developing the asset pricing model. Because that model is based on capitalization of rents, we then move to an in-depth study of rent determination, considering Ricardian and Thunen models and the modern but Thunen-derived bid-rent function approach. After briefly considering the macroeconomic influences on land markets, we return to the asset pricing model and analyze the reaction of land prices to increasing rents, growing demands from alternative land uses, and inflation. The possibility of capital gains from land ownership is examined, drawing a sharp distinction between windfall capital gains that result from events unexpected by most all market participants and routine, continuing capital gains for which an investor with no especial clairvoyance could plan.

In a concluding section, we use concepts from land market theory to briefly explore the implications of land use planning and regulation, public lands and wilderness protection, the new communication technologies and their impact on locational choice, value capture and public finance, and macroeconomic influences in land markets. Finally, we return to our opening theme: the role of land in the development of economics.

Despite our narrower focus on land markets, we draw upon all of the intellectual influences identified in the opening section. If any of these bodies of thought is slighted a little through underutilization, it is perhaps modern natural resource and environmental economics. Our relative inattention to that subject-matter will be forgiven, we hope, since the other chapters in this volume pay respectful attention to most all of its dimensions.

2. Theories of land rent and asset value

We begin with the simplest version of the asset pricing model. Land is treated as a marketable asset, valued only for its usefulness as a factor for producing a time stream of goods (e.g. agricultural commodities). The central role of rents in the asset pricing model and the "factor of production" perspective of land direct us to an examination of the Ricardian and Thunen models of rent determination. Then, we introduce the residential demand for land (which thus becomes a consumption good), and the relationships between land markets and financial markets. At that point, we return to the asset pricing model, in a more complete formulation, and examine the behavior of land markets in relation to other financial markets and in response to macroeconomic influences.

2.1. Capital value of land as a productive and marketable asset: A simple model

Assume that land is used only for producing marketable goods z subject to the production function $f(z, x, h, L)$. Then the rent accruing to an acre of land, p_h, is defined as $p_h = (\boldsymbol{p}_z \cdot z - \boldsymbol{p}_x \cdot \boldsymbol{x} - wL)/h$ in the timeless context. Introducing time, the present value of an indefinitely long stream of rents starting in $t = 0$ is:

$$W_0 = \int_0^{\infty} \mathrm{e}^{-rt} [p_h(t)] \mathrm{d}t. \tag{1}$$

Thus, if the land were sold at $t = 0$ and $p_h(t)$ represented the anticipated stream of rents accruing at all subsequent times, its market value as a productive asset would be $p_h^0 = W_0$. (Table 13.1 gives a full list of notation.)

If this land were owned by a producer who planned to sell it at some future time, T, its present value to its current owner would be:

$$W_0' = \int_0^{T} \mathrm{e}^{-rt} [p_h(t)] \mathrm{d}t + P_h^T \mathrm{e}^{-rT}. \tag{2}$$

However, P_h^T must surely be equal to the present value at T of the anticipated stream of subsequent rents, i.e.

$$P_h^T = \int_T^{\infty} \mathrm{e}^{-r(t-T)} [p_h(t)] \mathrm{d}t. \tag{3}$$

Thus, $W_0 = W_0'$. Starting at $t = 0$, the strategy of holding land for a finite time period, accruing rents during that period and selling the land at the end, is no more (and no less) rewarding than the strategy of holding it for an indefinitely long sequence of production periods. Given an unchanging set of expectations, the price of the asset at any time fully reflects the discounted value of future production. Thus, in the absence of clairvoyance, farming in anticipation of eventual capital gains from selling the land is no more (and no less) rewarding than farming in anticipation of continuing to farm.

Table 13.1
Notation

$U(\cdot)$	The utility function
$V(\cdot)$	The indirect utility function
$I[\cdot \mid \cdot]$	The income compensation function
$K = \{1, 2, \dots, k\}$	The index set for classes of consumers
y_k	The income of consumer k
z	The amount of the composite good
$\mathbf{z} = (z_1, \dots, z_i, \dots, z_n)$	The amount of commodities $1, \dots, i, \dots, n$
$N = \{1, 2, \dots, n\}$	The index set for classes of commodities
$z_i = a_i f(\cdot)$	The production function for commodity i
a_i	The propositionality factor that shifts the production function for commodity i
$\mathbf{x} = (x_1, \dots, x_j, \dots, x_m)$	The amount of inputs other than land and labor
h	The amount of land occupied
L	The amount of labor
z_i	Output per acre of commodity i
$\mathscr{L}$	Labor used per acre
p_z	Price of the composite good
p_{z_i}	Price of z_i
p_{x_j}	Price of x_j
w	The wage rate
F	Fertility of land
$F_{\min}$	Fertility of the poorest land in use
D	Distance (unless otherwise specified, from the central business district, CBD)
$s_{z_i}(D)$	Unit transportation costs of z_i to CBD
$s_{x_j}(D)$	Unit transportation costs of x_j from CBD
$,\dots,04s(D)$	Commuting costs to CBD
$c(D)$	Commuting time to CBD
$\mathbf{q}(D)$	The amount of amenities at D
$\tilde{L}$	Hours worked per day
l	Leisure hours per day
π	Profits per acre
p_h	Unit land rent
g	The growth rate in rent
$\hat{p}_h(D)$	The bid-rent function
$R(D)$	The rent gradient
$t = 0, 1, \dots, T, \dots, \infty$	Time in years
W_t	Net present value at time t of a revenue stream
P_h^t	The real market price of land at time t
NP_h^t	The nominal market price of land at time t
r	The real interest rate
a	The anticipated inflation rate
k	The proportion of initial land price borrowed, $0 \leq k \leq 1$
θ	The (flat) tax rate on current income, $0 \leq \theta \leq 1$.
ϕ	The (flat) tax rate on capital gains, $0 \leq \phi \leq 1$
$\rho = (1 - \theta)(r + a)$	The real nominal after-tax discount rate
β	The growth rate of nominal land price

Furthermore, if one assumes that land markets are fully integrated with financial markets, an entrepreneur without benefit of clairvoyance can anticipate normal returns from purchasing land for production and, perhaps, eventual resale. Abnormally high returns cannot, however, be anticipated.

The asset pricing model concludes that the market price of land is determined by the anticipated stream of rents it will produce and the returns from alternative investment opportunities, this latter consideration being reflected in the discount rate.

Now, we turn to the determination of rents.

2.2. Ricardian models and land quality

Assume land is available in effectively unlimited quantities. However, it comes in various qualities, so that the highest qualities of land are (at least, potentially) scarce.

Express land quality in a single dimension F, $0 \leq F < \infty$, where larger values for F indicate higher levels of soil fertility. Ignore non-land capital inputs and assume distance and location are of no importance. The problem is to determine how F influences the crops grown, the intensity of labor use and the rent accruing to land.

First, assume a single crop. Output, z_1, is a function of labor, land and soil fertility:

$$z_1 = f(L, h, F). \tag{4}$$

In general, $\partial^2 f/\partial L \cdot \partial F > 0$ and $\partial^2 f/\partial h \cdot \partial F > 0$, i.e. increasing fertility has a positive influence on the marginal productivities of both labor and land. Without invoking same specific restrictions on the production function, we cannot infer the sign of $\partial(L/h)/\partial F$, i.e. the influence of fertility on intensity of labor use per acre of land. Our first result is perhaps surprising: it is not generally true (nor generally untrue) that labor is used more intensively on the more fertile land.

Now, express the homogeneous production function in per acre terms, so that $\mathscr{z}_1 = z_1/h$ and $\mathscr{L} = L/h$:

$$\mathscr{z}_1 = a_1 f(\mathscr{L}, F), \qquad \partial f/\partial \mathscr{L} > 0, \quad \partial^2 f/\partial^2 \mathscr{L} < 0, \tag{5}$$

where a_1 is a proportionality factor ($a_i > 0$ for all i) characterizing the particular crop grown (this factor acquires more relevance in the multicrop case, below). Per acre profits, π, may be expressed as

$$\pi = p_{z_1} \cdot a_1 f(\mathscr{L}, F) - w\mathscr{L} - p_h(F). \tag{6}$$

At equilibrium, profits will everywhere be driven to zero. Thus, when labor use per acre is optimized at the level $\mathscr{L}^*$,

$$p_h(F)=p_{z_1}\cdot a_1 f(\mathscr{L}^*, F)-w\mathscr{L}^*. \tag{7}$$

Restricting $p_h(F)$ to be non-negative, there is some $F_{\min}$ for which $p_h(F_{\min})=0$ and $p_{z_1}\cdot a_1 f(\mathscr{L}^*, F_{\min})=w\mathscr{L}^*$. For $F>F_{\min}$, $p_h(F)>0$ and $p_{z_1}\cdot a_1 f(\mathscr{L}^*, F)>w\mathscr{L}^*$. Partially differentiating (7) with respect to F, we obtain $\partial p_h/\partial F=p_{z_1}\cdot a_1\partial f/\partial F$. Since $\partial f/\partial F>0$, it follows that $\partial p_h/\partial F>0$.

Our second result is that land rents are positive and increase with soil fertility for $F>F_{\min}$. At $F=F_{\min}$, land rents are zero. At $F<F_{\min}$, revenue is less than labor cost at the optimal factor combination. So, land with $F<F_{\min}$ is abandoned and output there is zero. If demand for z_1 was to shift rightward, driving p_{z_1} up, it is clear that a new solution, $F'_{\min}$, would be found for the zero rent fertility level, that $F'_{\min}<F_{\min}$, and that $p'_h(F)>p_h(F)$ for all $F>F'_{\min}$. Increased demand for crops results in cultivation of lower quality land at the margin and increased rents for all super-marginal land.

Now, consider the multicrop case. For crops in N, we have rent functions, analogous to (7) and differing only with respect to the proportionality factor:

$$p_{h_i}(F)=p_{z_i}\cdot a_i f(\mathscr{L}_i, F)-w\mathscr{L}_i, \qquad i\in N. \tag{8}$$

Optimizing with respect to labor use per acre, we obtain:

$$p_{z_1}\cdot a_1\partial f/\partial\mathscr{L}_1=w=p_{z_i}\cdot a_i\partial f/\partial\mathscr{L}_i, \quad \forall i\in N\setminus\{1\}. \tag{9}$$

Thus, for crops 1 and n,

$$\frac{p_{z_1}\cdot a_1}{p_{z_n}\cdot a_n}=\frac{\partial f/\partial\mathscr{L}_n}{\partial f/\partial\mathscr{L}_1}. \tag{10}$$

For any crop i, $p_{z_i}\cdot a_i$ is proportional to average value of product per acre. We seek to determine the relationship between value of crop, land use and land rent. Take the case where $p_{z_1}\cdot a_1>p_{z_n}\cdot a_n$. Eqs. (8) require that if $\mathscr{L}_1=\mathscr{L}_n$, $p_{h_1}>p_{h_n}$. Furthermore, if $\mathscr{L}_1>\mathscr{L}_n$, p_{h_1} would exceed p_{h_n} by no lesser amount than if $\mathscr{L}_1=\mathscr{L}_n$; an optimizing agent would not use additional labor with crop 1 unless revenue was increased by more than the additional labor cost. Thus, we know that when $p_{z_1}\cdot a_1>p_{z_n}\cdot a_n$ and $\mathscr{L}_1\geq\mathscr{L}_n$, $p_{h_1}>p_{h_n}$ for any $F>F_{\min}$.

Now, eq. (10) can be used to show that $p_{z_1}\cdot a_1>p_{z_n}\cdot a_n$ implies $\mathscr{L}_1>\mathscr{L}_n$. When $p_{z_1}\cdot a_1>p_{z_n}\cdot a_n$, $\partial f/\partial\mathscr{L}_n>\partial f/\partial\mathscr{L}_1$; which implies (given diminishing marginal productivity of labor) that $\mathscr{L}_1>\mathscr{L}_n$. Thus, $p_{h_1}>p_{h_n}$ for any level of soil fertility $F>F_{\min}$. Rent maximization would lead to a monoculture of the crop with the highest $p_{z_i}\cdot a_i$ $(i=1,2,\ldots,n)$ throughout the region. If, perchance,

$$\frac{p_{z_1}\cdot a_1}{p_{z_n}\cdot a_n}=1=\frac{\partial f/\partial\mathscr{L}_n}{\partial f/\partial\mathscr{L}_1}, \tag{11}$$

$\mathscr{L}_1 = \mathscr{L}_n$ and $p_{h_1} = p_{h_i}$. Thus, when $p_{z_1}/p_{z_n} = a_n/a_1$, producers would be universally indifferent as to which crop was produced on which land so long as $F \geq F_{\min}$.

These findings for the multicrop case depend on some special features of the specified relationship between production and soil fertility. Most crucially, we assumed that land quality can be characterized by a single, one-dimensional attribute F. If, more realistically, there were several dimensions of land quality and crops differed in the shapes of their isoquants among these dimensions, one would expect the regions to be divided into several monocultural zones of different crops. In addition there may be zones of indifference among two or more crops.

Until this point, we have (by treating all p_{z_i} as parametric) assumed that while people and land are immobile, goods may move costlessly into or out of the region. Imagine, instead, that the region were completely divorced from the outside world[8] and several goods were necessities (i.e. positive quantities were demanded at very high prices). Then, equilibrium prices $p_{z_i}^*$, would be established satisfying

$$\frac{p_{z_1} \cdot a_1}{p_{z_i} \cdot a_i} = \frac{\partial f/\partial \mathscr{L}_i}{\partial f/\partial \mathscr{L}_1}, \quad \forall\, i \in N \setminus \{1\}. \tag{12}$$

Positive amounts of the various crops would be grown, the exact quantities satisfying $p_{z_1}^*/p_{z_i}^* = a_i/a_1$, and producers would be universally indifferent as to which crop was grown on which land, so long as $F \geq F_{\min}$.

To summarize, highly simplified models in the Ricardian tradition yield the following results.

(1) When land is not scarce in quantity but is distinguished by a quality gradient with high quality land scarce: (a) there is some quality level $F_{\min}$ such that land with lower levels of F remains unused; (b) land with $F > F_{\min}$ attracts a positive rent, and the rent increases with F; and (c) when the price of output rises, some poorer land (the best of what was remaining) is brought into use, and the rent for all super-marginal land rises.

(2) It is unclear whether labor use per acre is more or less intensive on higher quality land.

(3) When the choice is among several crops, land quality (so long as it is defined on a unidimensional gradient) does not determine land use. With the economy open in goods, the general result is a monoculture of the crop with the highest per acre value of output. Note that this result depends on the assumption that the production functions for crops differ only with respect to the proportionality factor, a_i. If the economy is completely closed, commodity prices and quantities would be determined endogenously to satisfy (12). Nevertheless, pro-

---

[8] Samuelson (1959a) implicitly analyzes the closed economy case throughout.

ducers would be universally indifferent as to which crop was grown on which land, so long as land quality is at least $F_{\min}$.

2.3. *Thunen models and location*

Imagine a large city in the midst of a featureless, fertile plain, of uniform soil quality. At a great distance from the city, there is an impenetrable wilderness separating this state from the rest of the world. The question is: Assuming rational decision-making, how will distance to the city affect the choice of agricultural products and the intensity of land use? This, paraphrasing only slightly, is the problem posed by Thunen early in the nineteenth century. It is quite different from Ricardo's problem: Ricardo was concerned with the differential fertility of land. Thunen's concern was entirely with space and distance.

If some essential activity (e.g. market exchange) takes place at some given point, while activities (e.g. farming) occupy space so that no two farms can be on the same land, along any ray emanating from the market city one farm must be at a greater distance from the city than another. If distance imposes a cost (e.g. an increased expense for getting goods to market), this fact, together with the rational, profit maximizing behavior of farmers, is a sufficient basis for predicting the emergence of quite definite and in no way haphazard patterns of land use.

Given the location of the city, the question is: How does distance from the city affect which commodities are produced, the output of each, the intensity of labor use, and the rent accruing to land? First, consider a single crop. Express the neoclassical production function[9]

$$z_1 = a_1 f(\mathscr{L}), \qquad \partial f/\partial\mathscr{L} > 0, \quad \partial^2 f/\partial^2\mathscr{L} < 0. \tag{13}$$

While D, distance from the city, does not influence directly the production process, it does influence profits:

$$\pi = (p_{z_1} - s_{z_1}D)a_1 f(\mathscr{L}) - w\mathscr{L} - p_h(\mathscr{L};\ D). \tag{14}$$

At the zero-profit equilibrium, land rent is equal to revenue (net of transportation costs) minus labor cost:

$$p_h(\mathscr{L};\ D) = (p_{z_1} - s_{z_1}D)a_1 f(\mathscr{L}) - w\mathscr{L}. \tag{15}$$

Optimizing labor use requires:

$$\partial p_h/\partial\mathscr{L} = 0 = (p_{z_1} - s_{z_1}D)a_1 \partial f/\partial\mathscr{L} - w, \tag{16}$$

---

[9] Beckmann (1972) shows that the standard Thunen can be obtained with neoclassical production functions and, thus, are not dependent on the assumption that inputs are used in fixed proportions.

which implies:

$$\mathscr{L} = \left(\frac{\partial f}{\partial \mathscr{L}}\right)^{-1}\left(\frac{w}{a_1(p_{z_1} - s_{z_1}D)}\right). \tag{17}$$

Implicit differentiation of (16) yields:

$$\frac{\mathrm{d}\mathscr{L}}{\mathrm{d}D} = -\frac{\dfrac{\partial^2 p_h}{\partial \mathscr{L}\partial D}}{\partial^2 p_h/\partial^2 \mathscr{L}} = \frac{s_{z_1}(\partial f/\partial \mathscr{L})}{\partial^2 p_h/\partial^2 \mathscr{L}} < 0, \tag{18}$$

since $\partial f/\partial \mathscr{L} > 0$ and the second order conditions for profit maximization require $\partial^2 p_h/\partial^2 \mathscr{L} < 0$. Thus, labor use declines with distance. By (17), we find that labor use diminishes to zero at

$$\bar{D} = p_{z_1}/s_{z_1}, \tag{19}$$

which implies that land beyond $\bar{D}$ is abandoned.

It follows, since $\partial f/\partial \mathscr{L} > 0$ and $z_1 = 0$ when $\mathscr{L} = 0$, that output z_1, decreases with distance from the city and falls to zero at the critical distance $\bar{D}$. Furthermore, from (15) we see that land rent is zero when $\mathscr{L}$ and z_1 are zero at and beyond $\bar{D}$. Differentiating (15) with respect to D, we see that $\partial p_h/\partial D$ is directly proportional to $(\partial f/\partial \mathscr{L})(\partial \mathscr{L}/\partial D)$, which is negative. Thus, land rent decreases with distance.

Let us summarize the results for the single output case. Employment, output and land rent (all in per acre terms) are positive near the city and diminish with distance, reaching zero at the critical distance $\bar{D}$ at which unit transportation costs, $s_{z_1}D$, equal product price. If the product price were to rise relative to transportation costs, some more distant land would be brought into production, increasing total output and raising land rents and the intensity of labor use at locations nearer the city.

Now, we consider the multiproduct case. For simplicity, consider two commodities, $i = 1, 2$, with the rent functions:

$$\begin{aligned} p_{h_1}(\mathscr{L}_1;\ D) &= (p_{z_1} - s_{z_1}D)a_1 f(\mathscr{L}_1) - w\mathscr{L}_1, \\ p_{h_2}(\mathscr{L}_2;\ D) &= (p_{z_2} - s_{z_2}D)a_2 f(\mathscr{L}_2) - w\mathscr{L}_2. \end{aligned} \tag{20}$$

Optimizing with respect to labor use per acre, we obtain:

$$(p_{z_1} - s_{z_1}D)a_1 \partial f/\partial \mathscr{L}_1 = w = (p_{z_2} - s_{z_2}D)a_2 \partial f/\partial \mathscr{L}_2, \tag{21}$$

which implies:

$$\frac{(p_{z_1} - s_{z_1}D)a_1}{(p_{z_2} - s_{z_2}D)a_2} = \frac{\partial f/\partial \mathscr{L}_2}{\partial f/\partial \mathscr{L}_1}. \tag{22}$$

If $(p_{z_1} - s_{z_1}D)a_1 > (p_{z_2} - s_{z_2}D)a_2$ we know (by an argument analogous to that used in our analysis of the multicrop Ricardian model) that a monoculture of commodity 1 will be produced. Similarly, if the inequality is reversed, commodity 2 will be produced in monoculture. Assume, however, that commodities 1 and 2 are necessities and the regional economy is closed. Then, prices will adjust to ensure that both are produed, each in a separate zone of monoculture. The zones will be separated by a sharp boundary at D', which is determined so as to satisfy:

$$\left(p_{z_1} - s_{z_1}D'\right)a_1 = \left(p_{z_2} - s_{z_2}D'\right)a_2. \tag{23}$$

Which commodity is produced in the zone $0 < D \leq D'$ and which is produced in the zone $D' \leq D \leq \bar{D}$? Assume $s_{z_1} = s_{z_2}$ and $a_1 > a_2$ (i.e. that the ton-mile cost of transportation is the same for each commodity and the weight of commodity 1 produced with a given amount of labor is greater than that of commodity 2). Then as D increases, $s_{z_1}Da_1$ increases faster than $s_{z_2}Da_2$ and $(p_{z_1} - s_{z_1}D)a_1$ falls relative to $(p_{z_2} - s_{z_2}D)a_2$. Therefore, $(p_{z_1} - s_{z_1}D)a_1 > (p_{z_2} - s_{z_2}D)a_2$ for smaller values of D, and the inequality is reversed for larger values of D. Commodity 1, which weighs more per acre, is produced as a monoculture in the zone $0 < D \leq D'$, while commodity 2 is produced exclusively in the zone $D' \leq D \leq \bar{D}$. This result is demonstrated in Figure 13.1, where the rent function for product 1, $p_{h_1}(D)$, is more steeply sloped than that for product 2.

For more than two commodities, the logic of our analysis still holds. Zone boundaries are sharp, the commodity zones are arrayed from the market center in order of decreasing weight of product, and [since, by (20), land rent as a function of D is of steeper negative slope when a_i is larger for all values of D] each commodity will have only one zone of production.

These findings can be modified by making different assumptions about the production functions.

(a) If we permitted different functional forms for different commodities, distinct monocultural zones would again result, but there might be several non-adjoining zones of the same commodity.

(b) If two commodities could be jointly produced with continuously variable proportions of outputs, discrete zones disappear and the more transportable commodity is continuously substituted for the less transportable one with increasing distance.

Finally, by (23) and (22) it follows that at D', $\partial f/\partial \mathscr{L}_1 = \partial f/\partial \mathscr{L}_2$ and $\mathscr{L}_1 = \mathscr{L}_2$. The same amount of labor is applied to each commodity at the zone boundary. Thus, the smooth decline of the labor/land ratio with increasing distance is not disrupted at the zone borders. Land rents are also equal at the zone borders, but the rent curve for the commodity with higher transportation cost per acre is steeper.

In summary, labor use declines continuously and smoothly with distance; land rents decline continuously but with kinks at the zone boundaries; and the weight

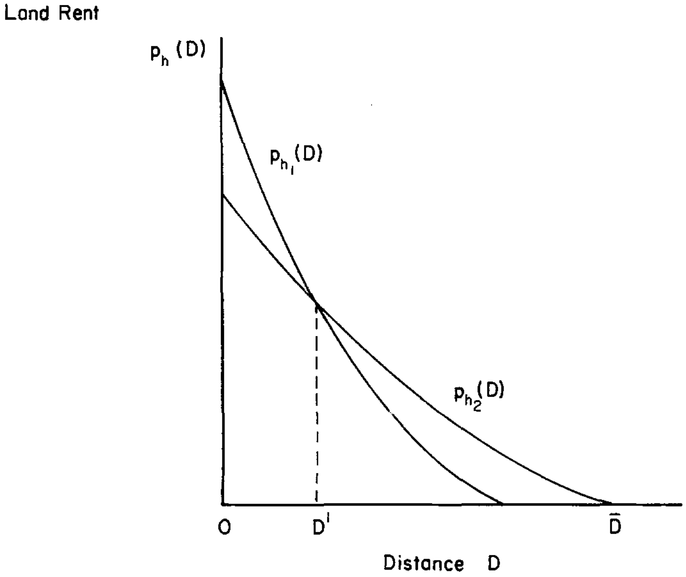

Figure 13.1. Assignment of commodities to Distance Zones. [Commodity 1 is produced exclusively in the zone defined by $0 < D \leq D'$, while 2 is produced exclusively in the zone $D' \leq D < \bar{D}$. Beyond $\bar{D}$, land is abandoned.]

of output per acre declines smoothly within zones but (most likely) with discontinuities at the zone boundaries. All of these results were obtained for models of linear space. If we had modeled Thunen's problem in two-dimensional space, the linear distance zones would be replaced by a series of rings concentric at the market place. While such a representation has more pictorial appeal, little is added in the way of definitive results.

Differences between Ricardian and Thunen results

In the Ricardian (R) models, one is concerned with land quality, F, but not with distance, D. In the Thunen (T) models, F is held constant and thus ignored while D is the primary focus. However, R models do not treat F in the same way that T models handle D. Output per acre is a function of F, but revenue per unit output is unaffected by F in R models. In T models, D does not appear in the production function. It does, however, appear in the revenue equation: unit revenue net of transportation cost decreases with increasing distance from the central market.

These different models obtain one consistent result. Land rents decrease with land quality until the poorest land is use brings no rent and all still poorer land is abandoned. Land rents decrease with increasing distance from the central market, until the most distant land in use brings no rent and all land more distant than that is abandoned.

In other respects, the R and T models generate quite different results, the T results being more interesting. R models obtain no general result about the relation between land quality and labor intensity; and they conclude that the choice of which commodity to produce is unrelated to (one-dimensional) land quality. T models find: different commodities are produced in each of a series of zones differentiated by distance from the market; within each zone there is a monoculture; commodities are assigned to zones systematically, and those with the greater weight of output per acre (for a given complement of labor) are assigned to zones nearer the market; labor intensity declines smoothly with increasing distance within and across zones; and, while land rent declines continuously with increasing distance, there are kinks at the zone boundaries.

2.4. *The bid-rent function approach*

The bid-rent function approach is now the standard treatment of residential location and urban housing markets. While the fundamental logic of this approach owes much to Thunen's work on the location of agricultural production – the organizing concept is the cost of distance from a central location – its implementation in a consumer choice context allows the analyst to take advantage of recent developments in demand theory: e.g. indirect utility, the expenditure function, and the duality results.

Consider a featureless plain with a central business district (CBD) at a given central location. All urban workers are employed in the CBD and must commute to work daily. Commuting cost is directly related to distance of the residence from the CBD. The consumer–worker enjoys land, h, and other consumption goods, z, and the budget constraint is effectively $y - s(D)$ (i.e. income minus commuting costs). The consumer–worker must choose an optimal residential location. Landlords, each of whom is a rent-maximizing local monopolist at a given location, adjust land rents to assure full occupancy. In this situation, what relationship emerges among distance from the CBD, land rent per acre, and residential density? The simplest bid-rent function analysis solves this urban land market problem. More complex models consider workers with different preferences, incomes and wage rates; multiple locations of interest (e.g. the CBD, a factory district, a port district, a suburban shopping center, a district with special amenities or attractions); and competition for land between urban and agricultural uses at the edge of the city.

2.4.1. A simple model of residential choice

Consider a consumer with utility function $U(D) = U[h(D), z(D)]$, facing a budget constraint $y - p_h(D)h(D) - p_z z(D) - s(D) = 0$. Utility is positively related to the amount of land (residential space) occupied and the quantity of the composite goods consumed. Rewriting the budget constraint, $p_h(D)h(D) - p_z z(D) = y - s(D)$, we focus on income net of commuting costs.

2.4.1.1. Diagrammatic analysis. The different allocations, A and B (Figure 13.2), provide the consumer with equal satisfaction, by construction. However, B is associated with a location at $D_2 > D_1$, and thus with higher commuting costs than A. Holding p_z, y and U^0 constant, we can determine the land rents, $p_h(D_1)$ and $p_h(D_2)$ which would make the consumer indifferent between the locations D_1 and D_2. At these rents, the more distant location involves the greater consumption of land and lower consumption of the composite good (Figure 13.2 and 13.3). From

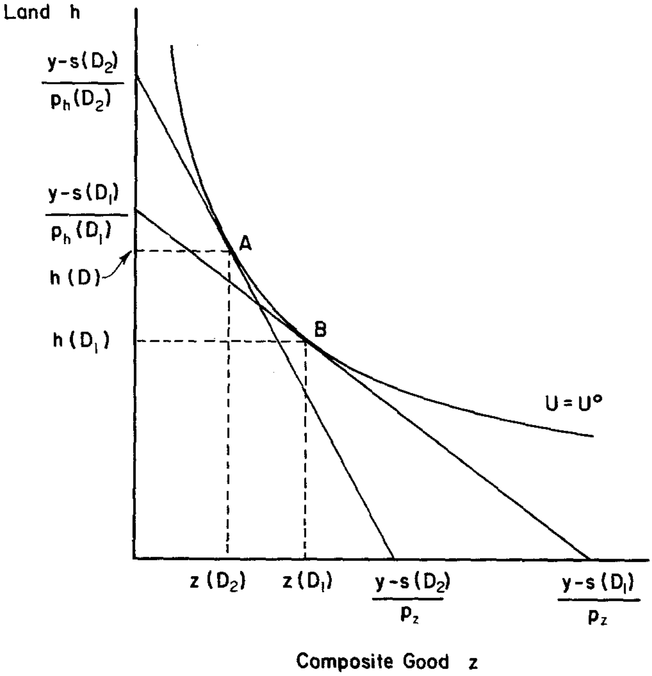

Figure 13.2. Consumer's budget allocation between land and other goods, given commuting costs that increase with distance from the CBD.

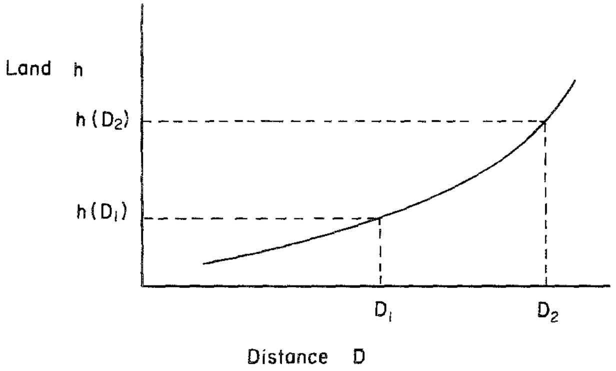

Figure 13.3. The relationship between land consumption and distance from the CBD.

Figure 13.2, we observe that:

$$\frac{p_z}{p_h(D_z)} > \frac{p_z}{p_h(D_1)},$$

which implies $p_h(D_1) > p_h(D_2)$. In order to hold utility constant, land rent must decrease with distance from the CBD (Figure 13.4).

Consider Alonso's (1964) definition of the bid-rent function: "... the set of prices for land the individual could pay at various distances while deriving a constant level of utility". It is evident that we have derived diagrammatically (Figures 13.2, 13.3 and 13.4) the individual consumer's bid-rent function, which is

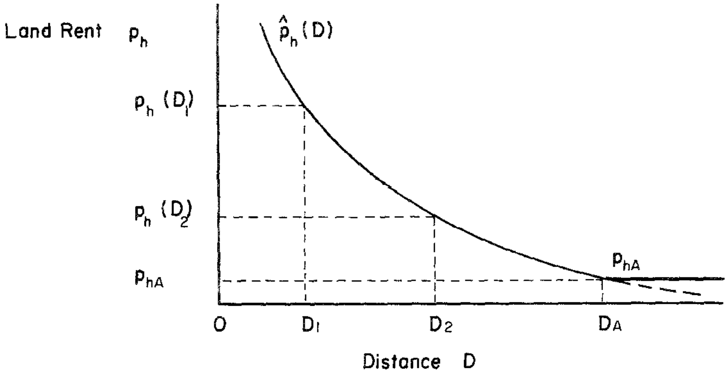

Figure 13.4. The bid-rent function, $\hat{p}_h(D)$, for urban workers and the agricultural rent p_{hA}.

defined as:

$$\hat{p}_h(D) = [y - s(D) - z(U^0, h)]/h. \tag{24}$$

The landlords are conceptualized as seeking to parcel their land into optimal-sized lots to maximize the rents they receive. Each is a local monopolist: one and only one parcel of land is located exactly at the specific distance D_i. The landlord's problem is to correctly anticipate the consumer's bid-rent function: i.e. the landlord i's rent asking behavior depends on D_i (which is given) and his estimate of the relevant demander characteristics [Anas (1982, p. 22)].

If all consumers are identical and each can move her place of residence costlessly, the bidding process will reach a static equilibrium in which the bid-rent function (Figure 13.4) becomes the rent gradient for the city. All urban workers will be indifferent as to where they live within the city, and each will achieve the same utility level as all others. At distances beyond D_A, land is used for agricultural purposes which yield a per acre rent of p_{hA} regardless of distance (by assumption, as is common in models of the urban land market).

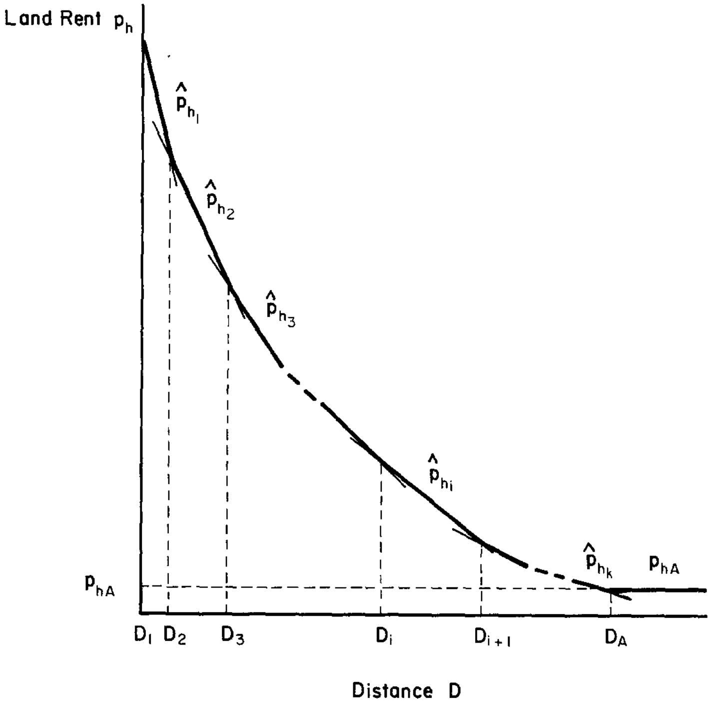

Figure 13.5. The rent gradient for a city with k classes of consumers, surrounded by farmland.

If there were k classes of consumers, $1, 2, \ldots, k$, with consumers being identical within classes and different across them, the rent gradient would be the envelope of the k different bid-rent functions (Figure 13.5). Consumers in class i will bid more than those in all other classes for land located between D_i and D_{i+1}, where their bid-rent function provides a segment of the rent gradient. Elsewhere, they will be outbid by members of some other class.

Faced with this rent gradient, consumers will exhibit the same choice behavior that generated it. Each will locate in the distance-zone where her bid-rent exceeds that of the members of other classes. Thus, each parcel of land is assigned to the highest bidder and none could gain by moving. In the fashion of Thunen agricultural land use models, consumers locate in k neighborhoods each homogeneous in terms of the class of consumers living there, with sharp neighborhood boundaries determined by the intersections of bid-rent functions.

2.4.1.2. Mathematical analysis. Now, we formalize the bid-rent function analysis. Define the consumer's indirect utility function [Gorman (1976), Varian (1978)] as:

$$\begin{aligned} V(D) &\equiv V[y - s(D), p_z, p_h(D)] \\ &\equiv \sup\{U[h(D), z(D)] : y - s(d) \geq p_h(D)h(d) + p_z z(D)\}. \end{aligned} \tag{25}$$

Using Roy's identity, the Marshallian demand functions are:

$$h^*(D) \equiv \frac{\partial V(\cdot)/\partial p_h(D)}{\partial V(\cdot)/\partial y} \equiv h^*[y - s(D), p_z - p_h(D)] > 0 \tag{26}$$

for land, and

$$z^*(D) \equiv \frac{\partial V(\cdot)/\partial p_z}{\partial V(\cdot)/\partial y} \equiv z^*[y - s(D), p_z, p_h(D)] > 0 \tag{27}$$

for the composite good.

Solving (25) for $p_h(D)$, we obtain the bid-rent function [Henderson (1977)]:

$$\hat{p}_h(D) = \hat{p}_h[V, y - s(D), p_z] \tag{28}$$

Substitute the bid-rent function into (25). Notice that:

$$V[y - s(D), p_z, \hat{p}_h(D)] = V = \text{constant}, \quad \forall D. \tag{29}$$

Now, take the total differential of (25) and hold utility, income and the price of the composite good constant:

$$\mathrm{d}V = \frac{\partial V}{\partial y}\,\mathrm{d}y + \frac{\partial V}{\partial p_z}\,\mathrm{d}p_z + \frac{\partial V}{\partial \hat{p}_h}\frac{\partial p_h(D)}{\partial D}\,\mathrm{d}D - \frac{\partial V}{\partial y}\frac{\mathrm{d}s(D)}{\mathrm{d}D}\,\mathrm{d}D;$$

$$\mathrm{d}V = \mathrm{d}y = \mathrm{d}p_z = 0 \tag{30}$$

implies:

$$\frac{\partial V}{\partial \hat{p}_h}\frac{\partial \hat{p}_h(D)}{\partial D} = \frac{\partial V}{\partial y}\frac{\mathrm{d}s}{\mathrm{d}D},$$

which implies:

$$\frac{\partial \hat{p}_h(D)}{\partial D}=\frac{\partial V/\partial y}{\partial V/\partial \hat{p}_h}\frac{\mathrm{d}s}{\mathrm{d}D}=-\left[h^*(D)\right]^{-1}\frac{\mathrm{d}s}{\mathrm{d}D}<0, \tag{31}$$

since $\mathrm{d}s/\mathrm{d}D>0$.

Thus, we have shown:

(1) The bid-rent function is decreasing with distance.

From the properties of the indirect utility function [Varian (1978)] it follows that:

$$\frac{\partial \hat{p}_h}{\partial V}=-\left[\frac{\partial V}{\partial p_h}\right]^{-1}>0, \tag{32}$$

$$\frac{\partial \hat{p}_h}{\partial y}=-\frac{\partial V/\partial y}{\partial V/\partial \hat{p}_h}>0, \tag{33}$$

and

$$\frac{\partial p_h}{\partial p_z}=-\frac{\partial V/\partial p_z}{\partial V/\partial \hat{p}_h}<0. \tag{34}$$

These results can be summarized:

(2) The bid-rent function is increasing in the utility level and income, and decreasing in the prices of non-land commodities.

Since (by assumption) $\partial h^*/\partial y>0$, it follows that

$$\frac{\partial h^*}{\partial s}=\frac{\partial h^*}{\partial y}\cdot\frac{\partial\left[y-s(D)\right]}{\partial s}=-\frac{\partial h^*}{\partial y}<0.$$

Examining the Slutsky equation, we would readily find that $\partial h^*/\partial p_z<0$.

Thus:

(3) If land is a non-inferior good, demand for land is decreasing in commuting costs and prices of consumption goods.

At this point, we turn to the expenditure function, which is defined by:

$$\begin{aligned} e(D) &\equiv e\left[p_z, p_h(D), U^0\right] \\ &\equiv \inf\left\{p_z z(D)+p_h(D)h(D): U\left[z(D), h(D)\right]\geq U^0\right\}. \end{aligned} \tag{35}$$

Introducing the bid-rent function, we have

$$e^0(D) \equiv e^0[p_z, \hat{p}_h(D), U^0]$$
$$\equiv \inf\{p_z z(D) + \hat{p}_h(D)h(D) : U[z(D), h(D)] \geq U^0\}, \tag{36}$$

and the following property holds:

$$y \equiv e^0[p_z, \hat{p}_h(D), U^0] + s(D), \quad \forall D. \tag{37}$$

Thus, if land rent follows the bid-rent function, the differences in minimum expenditures to maintain a given utility level are fully compensated by the differences in commuting costs as distance varies. Furthermore, the budget is everywhere exhausted.

The Hicksian compensated demand functions are:

$$\bar{h}(D) \equiv \frac{\partial e(\cdot)}{\partial p_h} \equiv \bar{h}[U^0, p_z, \hat{p}_h(D)] \tag{38}$$

for land, and for goods:

$$\bar{z}(D) \equiv \frac{\partial e(\cdot)}{\partial p_z} \equiv \bar{z}[U^0, p_z, \hat{p}_h(D)]. \tag{39}$$

From (37), we conclude:

(4) If land rent follows the bid-rent function, the Marshallian and Hicksian demand functions coincide. That is:

$$\bar{h}[U^0, p_z, \hat{p}_h(D)] \equiv h^*[y - s(D), p_z, \hat{p}_h(D)]$$

and

$$\bar{z}[U^0, p_z, \hat{p}_h(D)] \equiv z^*[y - s(D), p_z, \hat{p}_h(D)]. \tag{40}$$

Since the Hicksian demand curve is always downward sloping and $\partial \hat{p}_h / \partial D > 0$,

$$\frac{\partial \bar{h}(\cdot)}{\partial D} = \frac{\partial \left[\frac{\partial e^0(\cdot)}{\partial p_h} \right]}{\partial D} = \frac{\partial^2 e(\cdot)}{\partial^2 \hat{p}_h} = \frac{\partial p_h}{\partial D} > 0.$$

(5) The demand for land along the bid-rent function is increasing in distance.

Furthermore:

(6) If the commuting cost is constant per distance unit, the bid-rent function is strictly convex in distance.

To prove proposition 6, set $s(D) = a + bD$, where $a \geq 0$, $b > 0$ are constants. Then,

$$\frac{\partial p_h(D)}{\partial D} \equiv -[h^*(D)]^{-1}\frac{\mathrm{d}s}{\mathrm{d}D} = -b[\bar{h}(D)]^{-1}$$

from (31) and (40). Thus,

$$\frac{\partial^2 p_h(D)}{\partial^2 D} = -b\frac{\partial \bar{h}(\cdot)}{\partial D} < 0.$$

To this point, all results are for the individual consumer. Now, we consider the equilibrium for a city.

2.4.1.3. Equilibrium: Identical consumers. First, assume all consumers are identical in their preferences and income. Then, an equilibrium rent gradient must have the property that all consumers achieve the same utility level. If not, any consumer living at D_j and enjoying less utility than her counterpart at D_i would attempt to move to D_i, creating excess demand at that location. The usual adjustment processes would lead to an equilibrium in which all consumers achieve the same utility. An equilibrium rent gradient, $R(D)$, must therefore satisfy:

$$V[y - s(D), p_z, R(D)] = \text{constant}, \quad \forall D. \tag{29'}$$

That is to say, the rent gradient is identical to the bid-rent function of the typical consumer.

2.4.1.4. Equilibrium: Non-identical consumers. Now, consider a city with k consumers (or k classes of homogeneous consumers) each different from the others. Given the one-dimensional concept of distance we are using, each consumer i occupies an interval (D_i, D_{i+1}). Taking the rent of agricultural land as given, the boundaries between the zones occupied by different consumers are established through an adjustment process. In equilibrium, land is allocated everywhere to the highest bidder. Since the bid-rent function is decreasing in distance, it must be true that at a boundary D_{i+1},

$$\frac{\partial \hat{p}_{h_i}(D_{i+1})}{\partial D} < \frac{\partial \hat{p}_{h_{i+1}}(D_{i+1})}{\partial D}.$$

This implies:

(7) At each boundary between two worker-consumers, the inner consumer has a higher ratio of marginal commuting costs to land consumed than the outer consumer, i.e.

$$\frac{\mathrm{d}s_i(D_{i+1})/\mathrm{d}D}{h_i^*(D_{i+1})} > \frac{\mathrm{d}s_{i+1}(D_{i+1})/\mathrm{d}D}{h_{i+1}^*(D_{i+1})}, \quad i \in K \setminus \{k\}. \tag{41}$$

This result is often called the "Thunen condition", since the analogy is obvious with the condition that determines boundaries between zones of agricultural production.

Now, is there any way to determine which classes of consumers will occupy which distance zones? Assume all consumers have identical utility functions and commuting cost functions, but each of the k classes differs in income. If land is a normal or superior good, $\partial h^*/\partial y > 0$. Combining this with (41), we obtain:

(8) In the case of identical utility functions and commuting cost functions, if land is a normal or superior good those consumers with higher incomes will live further from the CBD, i.e. $y_i < y_{i+1}, i \in K \setminus \{k\}$.

However, if commuting takes time that could otherwise be used for work or leisure, this result should be reexamined. If we define $\tilde{L}$ as hours worked and l as leisure hours, the consumer's budget constraint becomes:

$$w\tilde{L} - p_h(D)h(D) - p_z z(D) - s(D) = 0,$$

where $24 - \tilde{L} - l - c(D) = 0$. At the optimal location, the marginal change in commuting costs with distance $(w[\mathrm{d}c(D)/\mathrm{d}D] + [\mathrm{d}s(D)/\mathrm{d}D])$ is equal to the marginal change in land costs $(-h[\mathrm{d}p_h(D)/\mathrm{d}D])$. The "Thunen conditions" become:

$$\frac{w_i\dfrac{\mathrm{d}c_i(D_{i+1})}{\mathrm{d}D} + \dfrac{\mathrm{d}s_i(D_{i+1})}{\mathrm{d}D}}{h_i^*(D_{i+1})} > \frac{w_{i+1}\dfrac{\mathrm{d}c_{i+1}(D_{i+1})}{\mathrm{d}D} + \dfrac{\mathrm{d}s_{i+1}(D_{i+1})}{\mathrm{d}D}}{h_{i+1}^*(D_{i+1})}, \quad i \in K \setminus \{k\}. \tag{42}$$

If income is derived from work, $y_i < y_{i+1}$ suggests that $w_i < w_{i+1}$. It is quite possible that, as income rises, the increasing real costs of commuting (out-of-pocket commuting costs plus the opportunity costs of commuting time) may dominate the tendency to increase land consumption. *A priori*, the income pattern of distance-zone occupancy is unclear. Observation suggests that the well-off tend to congregate in at least two zones: one of large residential estates near the edge of the city and one of fine townhouses near the CBD. To subsume this observation under standard residential location theory, however, it is necessary to introduce housing quality, amenities and (most likely) non-identical preferences. Models that consider all these influences can be constructed, but it seems essential to impose some rather arbitrary restrictions on them in order to obtain well-specified predictions. These restrictions limit the generality of the predictions obtained. On the other hand, researchers can incorporate *a priori* information to formulate well-specified conceptual models for empirical analysis of a wide variety of special cases (see Section 2.4.5 below).

2.4.2. Amenities and location

The amount of amenities at location D is expressed by the m-dimensional vector $\boldsymbol{q}(D)=[q_1(D),\ldots,q_s(D)]$. Utility can then be defined over the composite good, land and amenities:

$$V(D)=U[z(D),h(D),\boldsymbol{q}(D)]. \tag{43}$$

The indirect utility function is:

$$V(D)=V[p_z,p_h(D),\boldsymbol{q}(D),y-s(D)]. \tag{44}$$

Now, we introduce the income compensation function [Hurwicz and Uzawa (1971)], which represents the least amount of income the individual would require at location D to achieve the same utility as at location D^0:

$$\begin{aligned} &I[\boldsymbol{q}(D)\,|\,\boldsymbol{q}(D^0)]\\ &\quad=I[\boldsymbol{q}(D),p_h(D),s(D)\,|\,\boldsymbol{q}(D^0),p_z,p_h(D^0),y-s(D^0)]\\ &\quad=s(D)+\inf\{p_zz(D)+p_h(D)h(D):U[z(D),h(D),\boldsymbol{q}(D)]\geq V(D^0)\}, \end{aligned} \tag{45}$$

where

$$V(D^0)=V[(p_z,p_h(D^0),\boldsymbol{q}(D^0),y-s(D^0)]=V^0.$$

The bid-rent function can be determined by substituting $\hat{p}_h(D)$ for $p_h(D)$ in (45) and setting $I(\cdot)\equiv y$, for all D:

$$\hat{p}_h(D)=p_h[\boldsymbol{q}(D),y-s(D),V^0]. \tag{46}$$

In a city with identical consumers, $\hat{p}_h(D)=R(D)$, i.e. the individual bid-rent function coincides with the equilibrium rent gradient. Thus, the consumer's budget constraint can be stated: $y=p_zz(D)+\hat{p}_h(D)h(D)+s(D)$. Maximizing utility (43) subject to this constraint, we obtain the following first-order condition (among others):

$$\frac{\partial U/\partial q_i}{\partial U/\partial z}=\frac{1}{p_z}\frac{\partial\hat{p}_h}{\partial q_i}h,\qquad i=1,\ldots,s. \tag{47}$$

This indicates that, for a consumer to be in equilibrium, the marginal value of each amenity must be equal to the increment to land cost (i.e. the increment in unit rent multiplied by the area of land demanded).

The expression $(\partial\hat{p}_h/\partial q_i)h(\cdot)$ can be interpreted as the implicit price of amenity i [Diamond and Tolley (1982)]. This approach has been utilized to determine the benefits from land improvement projects and, more often, air pollution abatement [e.g. Freeman (1979), Polinsky and Shavell (1975, 1976), Polinsky and Rubinfeld (1977)].

2.4.3. The bid-rent function for producers

Return to Thunen's problem situation: A featureless plain with a central market where all trade takes place. In a competitive situation, land will be allocated to the highest bidding firms. Each firm is assumed to produce a single product z_i sold at the market, using a vector of inputs x purchased at the market and labor that resides on-site (e.g. on the farm). The product z_i is chosen from z. Each producer seeks to maximize land rents subject to a zero-profit constraint. Formally, for each z_i in z, maximize rent per acre:

$$\max p_{h_i}$$
s.t

$$\text{(a)}\quad \left[p_{z_i} - s_{z_i}(D)\right] z_i - wL_i - \sum_{j=1}^{m} \left[p_{x_j} - s_{x_j}(D)\right] x_{ji} - p_h \geq 0, \tag{48}$$

$$\text{(b)}\quad z_i = f(x_i, 1, L_i);\quad x_{ji} \geq 0;\ h = 1.\quad z_i = f(x_i, 1, L_i);\quad x_{ji} \geq 0.$$

If one substitutes the optimal values (L_i^*, x_i^*, and z_i^*) from the solution into (48a), which must always hold as an equality in equilibrium, and solves for p_h, one obtains:

$$\hat{p}_{h_i}(D) = \left(\left[p_{z_i} - s_{z_i}(D)\right] z_i^* - wL^* - \sum_{j=1}^{m} \left[p_{x_j} - s_{x_j}(D)\right] x_j^* \right), \tag{49}$$

and

$$\hat{p}_h(D) = \max\left[\hat{p}_{h_1}(D), \ldots, \hat{p}_{h_n}(D)\right]. \tag{49'}$$

As long as $\mathrm{d}s_{x_j}(D)/\mathrm{d}D > 0$ for all x and $\mathrm{d}s_{z_i}(D)/\mathrm{d}D > 0$, we obtain $\partial \hat{p}_h(D)/\mathrm{d}D < 0$, i.e. the bid-rent function is decreasing with distance. The emergence of monocultural product zones with sharp boundaries is assured. At the boundary between (say) crops i and k where $D_i \leq D_k$,

$$\frac{\mathrm{d}\hat{p}_{h_i}(D_i)}{\mathrm{d}D} < \frac{\mathrm{d}\hat{p}_{h_k}(D_i)}{\mathrm{d}D},$$

i.e. the bid-rent function for the firm producing the inner crop is steeper than that for the outer crop. This implies:

$$\left[\frac{\mathrm{d}s_{z_i}(D_i)}{\mathrm{d}D} z_i^* + \sum_{j=1}^{m} \frac{\mathrm{d}s_{x_j}(D_i)}{\mathrm{d}D} x_{ji}^*\right] > \left[\frac{\mathrm{d}s_{z_k}(D_i)}{\mathrm{d}D} z_k^* + \sum_{j=1}^{m} \frac{\mathrm{d}s_{x_j}(D_i)}{\mathrm{d}D} x_{jk}^*\right]. \tag{50}$$

This means that production processes with higher marginal transportation costs per land unit are located nearer the market, when marginal transportation costs include those for products and factors weighted by the optimal values for output and factor use. Eq. (50) generalizes similar results in Dunn (1957), Alonso (1964), and Miyao (1981, Theorem 1-1).

For completeness, we note that it is a fairly easy task to show, in the bid-rent function context as well as with a standard Thunen model, that labor use per unit of land decreases smoothly with distance from the market.

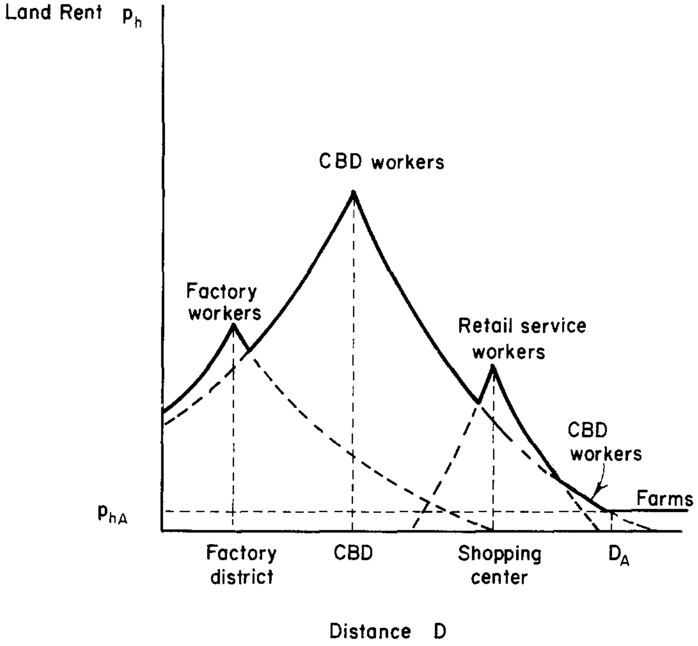

Figure 13.6. Rent gradient for a city with several workplaces.

2.4.4. *Location theory*

Both the Thunen and bid-rent function approaches provide starting points for a theory of land markets and location.[10] However, both approaches (as developed to this point in this essay) are confined to the simplest of cases: given a single point of interest (the market or the CBD) where do residents or firms locate? In this section, we briefly introduce the problems that arise in more complex – and more realistic – cases.

2.4.4.1. Firms and residents. If one assumes in advance the location of (say) a factory district, a CBD and a separate retail trade center, one can predict the residential location of various types of workers and the rent for land at various locations. Assuming identical utility functions of workers and no aesthetic distinctions among residential locations, a pattern such as that shown (Figure 13.6) would emerge.

---

[10] Major contributions to location theory include Lösch (1954), Izard (1956), Beckmann (1968), Greenhut (1970) and Takayama and Judge (1971).

2.4.4.2. Aesthetic distinctions, amenities and non-identical utility functions. If we admit aesthetic distinctions among residential locations, unequally distributed amenities, and non-identical utility functions, we make considerable progress toward realism, but at a non-trivial cost.

Realism is introduced by abandoning the "featureless plain" assumption. Cities are depicted in two dimensions rather than one. However, the distance-zones of the linear city are not automatically transformed into concentric rings. The well-to-do no longer congregate in a ring at the edge of the city. Silk stocking districts emerge (perhaps for mostly historical reasons; consider Philadelphia's main line). Expensive houses line aesthetically pleasing sections of ocean-front, lakefront or riverfront. Some of the well-to-do (weighted toward two-person households with both members employed) congregate in townhouses and condominiums near the CBD and attendant cultural attractions. Others, including many with growing families, live in fine houses on large lots in metropolitan-fringe suburbs with well-reputed school systems. Other socioeconomic classes also exhibit a greater diversity of residential location than the simple "concentric rings" models would predict.

In addition to aesthetically pleasing landscape features, neighborhood amenities include residential density and the age and quality of the housing stock, the characteristics of one's neighbors (including racial and ethnic composition, socioeconomic–sociodemographic characteristics, and propensity for crime and vandalism), and the level of services provided by local government. These neighborhood characteristics and others have been incorporated into models of residential choice (Diamond and Tolley).

The cost incurred by introducing these elements of realism is that location theory then tends to reduce to story-telling [Ward (1972)] and picture-drawing. The stories and pictures may be (and often are) pleasing and heuristically useful. However, the theory loses its power to produce general predictions: depending on the choice of givens and assumptions, almost anything is possible.

2.4.4.3. General equilibrium and endogenous determination of location. At a minimum, it seems, we must confront the location choices of firms and consumer–workers when both groups compete for sites in the urban area. However, major problems arise in so doing. Because employment is no longer restricted to predetermined locations, firm location and residential choice are both endogenous and interactive. Furthermore, there will be some advantages in establishing decentralized markets.

The simultaneous and endogenous determination of the existence and location of these multiple attractions (e.g. CBD, factory district, suburban shopping center) is a difficult problem. Dixit (1973) endogenously determined the location of the CBD, for the case of increasing returns. More generally, however, it seems the existence and location of the CBD, factory district, shopping center, etc. must be specified *a priori*. Then, one can ask the question: Given these things, where is my optimal location?

Koopmans and Beckmann (1957) published some troublesome results. Attempting to determine a rent maximizing location of production plants in a metropolitan area, they found that the conditions for the existence of an optimum sustainable by a price system are quite stringent. For example, if interplant transactions (i.e. intermediate goods) are considered, no sustainable optimum exists. Many subsequent authors have interpreted these and related findings as strongly suggesting the impossibility of general models in which all locational choices are endogenously determined. Game-theoretic formulations [see, for example, Found (1971), Harvey (1973), Vickerman (1980), Schweitzer (1983)], which recognize spatial interdependencies such that agent i's activities affect j's bid-rent function, tend to reinforce this negative conclusion.

An alternative approach involves the use of programming techniques to optimize various aspects of model cities. What is the optimal land rent gradient, the optimal population density, the optimal number of residents, etc.? Linear programming formulations abound[11] [Mills (1967), Mills and deFerrante (1971), Beckmann (1969), Solow and Vickrey (1971)], while others [e.g. Mirrlees (1972), Dixit (1973), and Riley (1973)] have used social welfare function formulations. These efforts have established some interesting general results, e.g. Mirrlees' (1972) finding that maximizing an egalitarian social welfare function may require that consumers (even if identical in preferences and incomes) living at different distances from the CBD achieve different levels of utility.

Nevertheless, these approaches are perhaps more useful for resolving specific empirical issues in urban planning, where the givens of the particular situation can be expressed as constraints. As such, they are helpful planning tools. However, their capacity to produce general results is restricted and – since the results obtained depend on the assumptions and constraints that are specified at the outset – it sometimes seems these models are themselves a rather sophisticated form of story-telling.

2.4.5. *Framework for empirical analysis*

While the search for general conclusions from realistically detailed formulations of the locational choice model is often frustrating, these models have provided a useful framework for empirical case studies. Variables of interest have been identified and fruitful hypotheses about relationships among variables have been specified.

---

[11] Frisch (1931), in a paper that analyzed important aspects of the Ricardian problem, used a specification that is an obvious and rather advanced precursor to linear programming.

The variables identified as relevant for explaining land rents comprise a considerable list. One needs first to consider the various sources of demand for land: residential, commercial, industrial, agricultural, forestry, minerals extraction, recreational, and ecosystem support. For those demands that are derived from the demand for raw materials, relevant variables include: resource quality, transportation costs, product demands, factor demand and supply relationships, production technologies, etc. Where the demand for land is derived directly from demands for services and amenities, relevant variables include distance-related costs, land quality (on-site, neighborhood and regional amenities, topography, etc.), factors influencing the supply and demand for complements (e.g. buildings), and demander characteristics including (but by no means limited to) wage rates and household composition. Finally, land use and rent will be determined by the outcome of competition between alternative sources of demand. Each land parcel will be assigned (in a competitive land market) to the use that generates the highest rent.

Important hypotheses have been specified for two kinds of relationships: those between behavior in locational choice and various explanatory variables; and those between land prices or rents and various characteristics of sites and the demander population.

Relationships of the first kind have been developed and empirically tested. For example, Smith (1982) investigated how racial preferences affect metropolitan housing markets and showed that such preferences may lead to segregated residential patterns even in the absence of present or past public policies of overt discrimination. Graves and Regulska (1982) estimated the effects on interregional migration of variation in amenities (climate, recreational opportunities, etc.) and sociodemographic variables (race and life-cycle variables). Pollard (1982) examined how view quality and access to the CBD jointly influence building height along the Chicago lakefront.

The second kind of relationships have often been established in the hedonic price analysis framework [Rosen (1974)] opening the door to estimating the benefits from improvements in on-site, neighborhood and regional amenities. There have been many hedonic analyses of benefits associated with quality of air [e.g. Harrison and Rubinfeld (1977)] and water [e.g. Brown and Pollakowski (1977)]. Vaughan and Huckins (1982) estimate the benefits from controlling urban expressway noise. Grimes (1982) estimates rent gradients connected with recreational amenities near a large urban center.

2.5. *Rents, asset markets and the possibility of capital gains in land*

In Section 2.1 we concluded that the capital value of land as a productive asset is equal to the net present value of the indefinitely long stream of rents it is anticipated to generate in the future. Furthermore, this is true at any time. At some future time T, the capital value of land is equal to the net present value of the rents anticipated from then onward. Thus, holding land in anticipation of a short stream of rents and a capital gain is no more (or no less) rewarding than holding it in anticipation of an indefinitely long stream of rents.

As eq. (2) indicates, given a set of unchanging expectations about interest rates and the time-stream of rents, the market price of land, p_h^t, at any future time, t, is predictable at the outset. The standard theory of rent determination has been reviewed in the intervening parts (Sections 2.2–2.4). That theory is basically microeconomic and general equilibrium in nature. However, it is widely believed that land prices are also influenced by events in the macroeconomy, taxation policy, and the institutions that govern financial markets. Thus, conventional discussions of land prices [e.g. Scott (1983)] make frequent reference to inflation, capital gains taxation, leverage, institutional rigidities that influence mortgage interest rates and the returns available to small savers from investments other than land. It seems clear that through the 1950s, 1960s and 1970s investments in land (including but not limited to farmland) outperformed most other broad categories of investments and much of the above-normal rewards to landowners came in the form of capital gains [Boyne (1964), Castle and Hoch (1982)]. Often, a substantial component of the capital gains in land is attributed, in conventional discussions, to macro and institutional considerations.

It is by no means otiose to ask about the relationship of land prices to macro and financial–institutional variables. Land is a productive asset, as the relationship of land prices to rents makes clear. It is also a real asset serviceable for wealth-holding (in this respect, it is like gold). Furthermore, its ready marketability makes it a financial asset. Being both a financial and real asset, land is a preferred form of security for loans and thus land ownership is an effective way of acquiring indebtedness. Perhaps these aspects of the land asset influence markets therein.

Our prime purpose in this section is to enquire into the possibility of capital gains in land. An interesting secondary question is whether the real-asset and financial-asset attributes of land modify the fundamental conclusions of the asset-pricing model, which focuses directly on the productive-asset nature of land, i.e. on the rents that accrue to land in reward for its production of goods and services.

2.5.1. *On the possibility of above-normal returns to land*

The total return to the landowner includes a stream of rents and a capital gain, the latter being the difference between the net sale price of the land and its purchase price. (Either or both of these components of total return may take positive, zero or negative values.) As eq. (2) indicates, the present value of land is the discounted value of the stream of rents through the time of sale and the net sale price. Assuming many potential buyers, the market price of land in the initial period will equilibrate at its present value. This implies that in equilibrium the market price of land will be established so that it will earn its purchaser a return equal to the return from the same amount of capital invested at the market rate of interest. That is, the market price for land will be established so that landownership yields a normal total return.

This analysis ignores all of the uncertainties that make prediction of future rents, sale prices and interest rates a difficult task. If it occurs that rents or interest rates begin, at some time $t' > 0$, to deviate from what was initially anticipated, the market price at t' will reflect this turn of events. An increase in rents or a drop in the real rate of interest will cause $p_h^{t'}$ to be higher than it would otherwise have been. An owner who had purchased at $t = 0$ would find, beginning at t', that she was earning an above-normal return based on her actual purchase price at $t = 0$. A decrease in rents or a rise in real interest rates would have the opposite effect.

These kinds of unanticipated events are one-shot changes in the anticipated stream of rents and in the land price path. A buyer purchasing after such events are fully anticipated would expect to earn normal returns from that time onward. All new information would have been reflected in the one-shot asset price adjustment.

Thus, we conclude that new information emerging during the period of ownership may permit an owner to enjoy above-normal returns, calculated on the basis of the initial purchase price. In such cases, the owner may be said to have enjoyed a windfall gain. If, as we believe, asset markets are rational and efficient in the sense that current asset prices at any moment fully reflect all available information, it is not generally possible to plan to enjoy above-normal returns from land ownership. In other words the generally expected returns at the current asset price cannot be above normal. One could gamble for above-normal returns, betting that one's information is superior to that which the market is reflecting. But in so doing, one is substituting one's own expectations for those that motivate the market.

In summary, changes in the expected total returns from land relative to other investments will be reflected, nearly instantaneously, in the market price of land. More optimistic expectations would lead to one-shot capital gains for established land owners. Following these adjustments, the total returns from land ownership will again be equilibrated with those from alternative investments.

Capital gains of this type cannot be attributed to trends (e.g. growth in population and income, or continuing inflation) or existing institutional factors (e.g. preferential taxation of capital gains, rigidities in financial institutions, or farm commodity programs). Rather, such capital gains can only arise from favorable changes in expectations about these things. Unfavorable changes leading to capital losses are, *a priori*, equally likely.

2.5.2. *On the possibility of anticipated capital gains*

Fully anticipated above-normal total returns to asset ownership are impossible, we have concluded, since they are inconsistent with efficient asset markets. Unanticipated gains and above-normal returns may, of course, occur. A more interesting question is whether fully anticipated capital gains can occur, within the context of normal total returns.

2.5.2.1. Growth in rents. Assume land rents are expected to grow indefinitely at the rate g, so that $p_{ht} = p_{h0}e^{gt}$. Then, eq. (1) can be rewritten:

$$P_h^0 = \int_0^\infty e^{-rt}\left(p_{h0}e^{gt}\right)\mathrm{d}t = \frac{p_{h0}}{r-g}. \tag{51}$$

So long as $r > g$ (and it is difficult to imagine an indefinitely long time sequence in which $r \leq g$), P_h^0 is bounded. Note that P_h^0 will be larger when g is larger. Again, current asset prices reflect the anticipated growth in rents, eliminating any general expectation of above-normal returns.

The growth in rents is reflected in the time path of asset prices:

$$P_h^t = \int_0^\infty e^{-r\tau} p_{ht} e^{g\tau}\mathrm{d}\tau = \frac{p_{ht}}{r-g} = \frac{p_{h0}e^{gt}}{r-g} = P_h^0 e^{gt}. \tag{52}$$

That is, the price of land will continue to grow at the rate g, exactly the same rate at which rents are growing.

Growth in rents implies positive capital gains (52) but also a higher initial price (51) that reduces total returns to the normal level. This higher initial price resulting directly from the growth in rents – and therefore the growth in land prices – is what Crowley (1974) and Castle and Hoch (1982) have called the capitalization of capital gains.

Moreover, the greater is g, the larger is P_h^0/p_{h0} (i.e. the ratio of asset price to rent in the initial period). Given that $P_h(t)$ and $p_h(t)$ grow at the same rate, higher rates of growth in rents imply that greater proportions of the total return from land ownership accrue as capital gains rather than current rents. When land rents are growing land is a "growth stock". Melichar (1979) suggests some of the policy implications for agriculture of a farm economy with high land prices, low rents in the early years and a high proportion of total returns coming in the form of capital gains or "deferred income".

2.5.2.2. Changes in land use. Thunen's location theory and its modern bid-rent successors predict that any piece of land will be assigned to the use that returns the greatest rent. For any plot of land there will be several possible uses, each offering some given rent. Only the highest rent use will be selected and the rents available from lower-rent uses will have no influence on the amount of rent received.

However, imagine that there is some lower-valued land use whose rent is growing at g' while the rent from the current higher-valued use is growing at g, and $g' > g$. Thus, it is predictable that, at some time, t^*, in the future, land use will change as the faster-growing rent stream intersects the initially higher but slower growing stream. Perhaps growing recreational demands will eventually displace forestry, or growing urban demands will displace farming. The interesting question is whether the yet-ineffective but faster-growing demand for land influences its price well before land use actually changes.

For example, denote the rent stream from agricultural uses as $p_{hf}(t)$ while that from urban uses is $p_{hu}(t)$. Assume $p_{hf0} > p_{hu0}$ but $p_{hf}(t)$ grows at the rate g while $p_{hu}(t)$ grows at $g' > g$. The time pattern of rents will be as follows: for $0 \leq t \leq t^*$, the rent is $p_{hf}(t)$, which grows at g; at t^*, $p_{hft^*} = p_{hut^*}$; for $t \geq t^*$ the rent is $p_{hu}(t)$ which grows at the faster rate g'.

It is clear from the earlier analysis that after t^* land prices will grow at the higher rate g', generating greater real capital gains. What is not immediately obvious is the behavior of land prices during $0 \leq t \leq t^*$.

The price of the land at $t = 0$ is the discounted value of the stream of agricultural rents for $0 \leq t \leq t^*$ plus the discounted value of the stream of urban rents thereafter. Thus,

$$\begin{aligned} P_h^0 &= \int_0^{t^*}\left(\mathrm{e}^{-rt}p_{hf0}\mathrm{e}^{gt}\mathrm{d}t\right) + \mathrm{e}^{-rt*}\left\{\int_0^{\infty}\mathrm{e}^{-rt}p_{hu0}\mathrm{e}^{g'(t^*+t)}\mathrm{d}t\right\} \\ &= p_{hf0}\int_0^{t^*}\mathrm{e}^{-(r-g)t}\mathrm{d}t + \mathrm{e}^{-rt*}p_{hu0}\mathrm{e}^{g't^*}\int_0^{\infty}\mathrm{e}^{-(r-g')t}\mathrm{d}t \\ &= p_{hf0}\frac{1}{r-g}\left(1-\mathrm{e}^{-(r-g)t^*}\right) + \left(p_{hu0}\mathrm{e}^{g't^*}\right)\mathrm{e}^{-rt*}\frac{1}{r-g}. \end{aligned}$$

Since $p_{hu0}\mathrm{e}^{g't^*} = p_{hut^*} = p_{hft^*} = p_{hf0}\mathrm{e}^{gt^*}$,

$$P_h^0 = \left(\frac{p_{hf0}}{r-g}\right)\left\{1 + \frac{g'-g}{r-g'}\mathrm{e}^{-(r-g)t^*}\right\}. \tag{53}$$

Examining the behavior of the time derivative of P_h^t as t goes from 0 to t^*, we find that the growth rate of P_h^t is always positive; approaches g as t approaches zero; and is equal to g' at t^*. Thus, the growth rate of P_h^t increases smoothly from a rate barely above g when land use conversion is still in the distant future to the rate of g' at the conversion date and thereafter. The prospect of future changes in land use and subsequent increases in rents, even though conversion may be many years in the future, generates a pattern of increasing real capital gains, small at first but gathering momentum as the conversion date approaches.

A special case with perhaps more heuristic appeal is that where $g = 0$, i.e. real rents from farming are not growing. In this case, the real price of land (which would be stable if there were no prospects of eventual conversion) will commence increasing, slowly at first, many years before conversion takes place. The growth rate of land prices will increase smoothly until it reaches g' at t^*, and remain at g' thereafter.

2.5.2.3. Changes in the general price level.

A major indicator of macroeconomic performance – and hence an important target variable for macroeconomic policy – is the general price level. Inflation is defined as an increase in the general price level, while deflation is a decrease therein.

The possibility of changes in the general level of prices leads to a distinction between nominal prices (which include the effects of inflation and/or deflation) and real prices (which are nominal prices corrected for such influences). It is obvious that changes in the general price level would be likely to influence the nominal price of land. A more interesting issue is whether changes in the general price level affect the real price of land.

While inflations and deflations are possible, and periods of both have occurred in this century,[12] the relative prevalence of inflationary periods has lead most economists to use the term, inflation, when referring to changes in the general price level. Using this terminology, a deflationary period is one in which the rate of inflation is negative.

If inflation at the rate $a > 0$ is anticipated, the nominal interest rate will be $r + a$. Assuming that land rents are affected by inflation just as is the general price level,[13] nominal land rents will increase at the rate $g + a$. The nominal price of land at $t = 0$ (i.e. NP_h^0), under these conditions, is:

$$NP_h^0 = \int_0^\infty \mathrm{e}^{-[(r+a)-(g+a)]t} p_{h0}\,\mathrm{d}t = \frac{p_{h0}}{r-g} = P_h^0. \tag{54}$$

Thus, anticipated inflation has no influence on the real or nominal price of land at $t = 0$.

However, with the passage of time, $NP_h(t)$ and $P_h(t)$ diverge. $NP_h^t = NP_h^0 \mathrm{e}^{(g+a)t}$, while $P_h^t = P_h^0 \mathrm{e}^{gt}$; that is, the real price of land grows at the rate g, while its nominal price grows at the rate $g + a$. If nominal land rents inflate at the same rate as the general price level, land ownership is a perfect hedge against anticipated inflation.[14] Land ownership will not, however, outperform other investments given a set of inflation expectations.

---

[12] The data series on farmland prices generated by Castle and Hock (1982) includes long periods of inflation, but also the shorter but sharp deflation of the early 1930s.

[13] This assumption permits us to abstract from changes in relative prices in order to examine the influence of changes in the general price level.

[14] Harris' (1979) finding to the contrary depends on his rather eccentric assumption that nominal land rents remain constant while the general price level inflates.

2.5.2.4. Inflation, leverage and capital gains taxation. Real estate is a preferred form of security for loans and, thus, ownership of land and improvements thereupon serves as a way of acquiring indebtedness. Land-secured indebtedness is desirable when land prices are expected to grow with time. In such circumstances, the smaller one's initial equity the greater the expected return on one's investment. This relationship is known as the principle of leverage.

During (fully anticipated) inflationary periods, the nominal price of land grows at the inflation rate. A debt-free land owner is protected from inflation but does not profit from it. On the other hand, an indebted land owner profits from inflation since his equity grows faster than the rate of inflation.

Capital gains are taxed, in many jurisdictions, more leniently than current income, thus encouraging the pursuit of income in the form of capital gains.

Conventional wisdom suggests that, during the inflationary decade beginning in the late 1960s, many middle-class Americans were able to acquire unaccustomed wealth via home or farm ownership, due to the combined effects of inflation, leverage and capital gains taxation. We now examine the combined effects of these phenomena on land prices.

Assume fully anticipated inflation at the rate $a > 0$, a flat income tax at the rate $\theta (0 \le \theta \le 1)$, a flat capital gains tax at the rate $\phi (0 \le \phi \le 1)$, nominal land prices grow at the rate β, and the landowner's initial equity is $(1-k)NP_h^0$ where $k(0 \le k \le 1)$ is the proportion of the purchase price borrowed. Introducing these considerations,[15] we obtain:

$$\begin{aligned} NP_h^0 = \sum_{i=0}^{\infty} \bigg\{ & \int_0^T \left[(1-\theta)\left(p_{h0} e^{(g+a)iT}\right) e^{(g+a)t} e^{-\rho t}\right] dt \\ & - \int_0^T \left[(1-\theta)\left(NP_h^0 e^{\beta iT}\right) k(r+a)\left(1-\frac{1}{T}t\right) e^{-\rho t}\right] dt \\ & + (1-\phi)\left(NP_h^0 e^{\beta(i+1)T} - NP_h^0 e^{\beta iT}\right) e^{-\rho T} \bigg\} e^{-\rho(iT)}, \end{aligned} \tag{55}$$

where $\rho = (1-\theta)(r+a)$ is the after-tax nominal discount rate, and $(1-\theta)(r+a)\rho^{-1} = 1$.

After some manipulations, this reduces to:

$$NP_h^0 = \frac{(1-\theta) p_{h0} (\rho - g - a)^{-1}}{\begin{aligned} 1 + \bigg\{ & \frac{(1-\theta)(r+a)}{\rho} k\left[(\rho T)^{-1}(e^{-\rho T} - 1) + 1\right] \\ & - (1-\phi)(e^{\beta T} - 1) e^{-\rho T} \bigg\} (1 - e^{(\beta-\rho)T})^{-1} \end{aligned}}. \tag{56}$$

---

[15] Lee and Rask (1976) consider many of these variables in their model to predict individual bids for farmland, given exogenously determined land prices and growth rates thereof and data on the individual's economic and financial circumstances. Our model is designed to determine endogenously the market equilibrium price of land and its growth rate, in the case of identical farmers.

Table 13.2

Example no.	ρ	β	T	k	ϕ	$(1-\phi)B(kA)^{-1}$
1	0.05	0.04	30	0.7	0.2	1.23
2	0.05	0.04	25	0.7	0.2	1.31
3	0.06	0.04	30	0.7	0.2	0.82

Note that the numerator of (56) is analogous to (54). Thus, the denominator of (56) is of interest. First, (56) is bounded, so long as $\beta < \rho$ and $g + a < \rho$, as they must be in the long run. Second, examination of (56) indicates that when $\phi < \theta$ (i.e. when capital gains are taxed at favorable rates), NP_h^0 is larger.

Now, rewrite (56) as follows:

$$NP_h^0 = \frac{(1-\theta)p_{h0}(\rho-g-a)^{-1}}{\{1+k[A]-(1-\phi)B\}(1-e^{(\beta-\rho)T})^{-1}}. \tag{56'}$$

Then, kA is always positive when $k > 0$ and increases when k is large, implying that the need to pay interest drives down the bid-price for land. B is always positive when $\beta > 0$, implying that nominal capital gains increase the nominal bid-price for land. The net effect of these influences on NP_h^0 is unclear, without recourse to empirical assumptions.

We examine the size of $(1-\phi)B(kA)^{-1}$. If it exceeds 1 (i.e. the capital gains effect exceeds the debt service effect), the net effect of both is to increase NP_h^0, the initial nominal price of land. In Table 13.2 we consider some numerical examples.

Observe, first, that when β is almost as large as ρ (i.e. the growth in nominal land prices is almost as great as the nominal after-tax discount rate), the capital gains effect dominates and NP_h^0 increases. Comparing examples 1 and 2, we see that a shorter loan term further increases NP_h^0. Second, comparing examples 1 and 3, we see that when ρ substantially exceeds β, the debt service effect dominates and NP_h^0 is reduced.

Assuming competitive equilibrium conditions in asset markets, we readily deduce that $\beta = g + a$, that is, the growth rate of nominal land prices is equal to the growth rate of nominal rents, as before. Leverage, capital gains taxation and anticipated inflation do not provide an explanation of unusually high continuing capital gains or continuing above-normal returns to landowners. They do, however, influence the initial price of land.

2.5.3. *Unanticipated capital gains (again)*

Equation (56) provides us the opportunity to consider the effect on NP_h^0 of a sudden change in expectations about the general price level or taxation rates.

Considering tax rates, if ϕ fell relative to ϕ (i.e. a change in tax rates favored capital gains relative to current income) land prices would rise, so long as $\beta = g + a > 0$.

If inflation expectations racheted upwards from a to $a' > a$, while fixed-interest-rate loans permitted an indebted landowner to continue paying nominal interest of $r + a$, (56′) would read

$$NP_h^{0'} = \frac{(1-\theta)p_{h0}\left(p' - g - a'\right)^{-1}}{1 + \left\{\dfrac{(1-\theta)(r+a)}{\rho'}k[A] - (1-\phi)B\right\}\left(1 - e^{(\beta-\rho')T}\right)^{-1}}, \tag{56''}$$

where $\rho' = (1-\theta)(r+a')$ and $(1-\theta)(r+a)\rho'^{-1} < 1$. By inspection, $NP_h^{0'} > NP_h^0$. Under these conditions unanticipated increases in inflation – or unanticipated institutional changes favorable to land prices – would further increase the nominal price of land, benefiting established landowners.

The above-normal returns to land ownership during the 1950s, 1960s, and 1970s [Boyne (1964), Castle and Hoch (1982)] are inconsistent with explanations based on economic trends anticipated well in advance. They are not, however, inconsistent with explanations that include a series of increases in inflation expectations.

The fairly sharp drop in midwestern farmland prices in the five years following 1978 [Scott (1983)] is not inconsistent with explanations that include reductions in expected grain prices and inflation rates. Additional explanation for the post-1978 drop in nominal land prices is provided by the increase in the real component of interest rates during that period. Comparison of examples 1 and 3 in Table 13.2 shows the strong land-price-depressing influence of an increase in ρ relative to β. This effect can be attributed to a relative increase in the r component of ρ compared to the g component of β (i.e. to the relative increase in real interest rates compared to the growth in rents).

As this analysis again demonstrates, it is essential in explaining land price movements to be aware of the distinction between predictable capital gains (and losses) and those that are windfalls resulting from changed expectations after initial purchase.

3. Extensions and implications

The theory of land rents and land prices that we have outlined treats land prices as the capitalized value of rents, and shows how rents are generated from the use of land in production and consumption. In so doing it offers explanations of land use, investment in improvements to land, the location of economic activity, and the behavior of markets in land assets. Land market theory has direct links with public finance, regional economics and applied welfare economics, to name just a few related subdisciplines.

Just as land markets are well integrated into the production and consumption sectors and the financial and asset markets, land market theory is well integrated into the modern economic theory of production, consumption, capital and finance. Land market theory exhibits many of the characteristics of modern mainstream economic theory. It derives much of its power from the consistent and tenacious application of a few strong behavioral and structural premises. Maximizing individual behavior, exclusive and rival goods and amenities, and well-functioning, rapidly-adjusting markets with many participants account for the precise predictions that emerge from land market theory and the optimality of the predicted outcomes. A theory with such a simple yet robust structure seems to invite an ever-expanding array of applications.

The simplicity of this theory (a source of robustness) leads to some inconsistencies with the observed reality of some important situations. This, of course, invites analyses that drop overly restrictive premises – for example, the assumption that all inputs, goods, services and amenities are strictly exclusive and rival – and examine the consequences.

In this concluding section, we briefly introduce some of the extensions and implications of land market theory. Our coverage is incomplete and merely suggestive, but hints at the richness and diversity of the possibilities that exist. We consider some fairly direct extensions of standard land market theory, some cases where its results must be tempered with an understanding that the situational realities sometimes diverge from the simple premises of the standard theory, and some linkages between land markets and the macroeconomy. Finally, we speculate about the importance of land and related concepts to the future development of economic theory.

3.1. *Land use*

As we have shown, land market theory uses a rather restrictive set of assumptions to predict land use and land prices and to demonstrate the optimality (or, more accurately, efficiency) of the predicted outcomes.

Nevertheless, a considerable array of allocative concerns about land use find public expression. Governments are petitioned to act upon, and on occasion respond favorably to, concerns about: incompatible land uses; the perceived need to provide for a continued stream of amenities (ranging from greenbelts and farmscapes to pristine wilderness) at public expense or through police power coercion of landowners; premature conversion of land to urban uses; and the preservation of the long-term productive potential of farm and forest land.

Allocative theorists may choose to reject these various claims of public concern as manifestations of self-interested rent-seeking behavior. Alternatively, it may be argued that non-exclusiveness and non-rivalry are endemic to land use issues and, accordingly, claims of the social of efficiency of land markets are without support. The relevance is obvious of non-exclusiveness and non-rivalry to the problems of incompatible land uses and failure of markets to provide the efficient level of collective amenities. Intertemporal market failure is a more difficult argument to make, but many authors have claimed that private and social discount rates diverge with the social rate being the lower, while Ferejohn and Page (1978) make a more fundamental attack on the ethics of discounting *per se*. These market failure arguments have been used in support of a wide range of public initiatives in land use planning and zoning, subdivision regulation, building codes, farmland preservation and soil conservation.

Economists have addressed the issues that arise when government attempts to ameliorate perceived inefficiencies in land markets using, for example, zoning regulations. The economics of the transition of land from rural to urban use has been examined with theoretical models [Arnott and Lewis (1979)] and econometric analyses of land sales[16] [Chicoine (1981)]. Zoning institutions have attracted the direct attention of economists. Some have attempted to identify optimal zoning policies [Helpman and Pines (1977) and Grieson and White (1981)]. Others have developed conceptual arguments and assembled empirical evidence to the effect that zoning institutions really make little difference to outcomes in the land market [Siegan (1970) and Maser, Riker and Rosett (1977)].

The "zoning makes no difference" argument is based on theories of endogenous government, and in its simplest form claims that the very same social and economic influences that shape land markets will be brought to bear on zoning institutions. Thus, it is claimed, zoning not so much directs the land markets as follows it.

Some land market theorists have challenged the need for the involvement of national government in securing amenity levels, with the following subtle argument originally attributable to Tiebout (1956). They claim that more market opportunities exist than are immediately apparent for individuals to effectively express their demands for amenities. In particular, individuals selecting a residential location exercise a considerable degree of choice about amenity levels. Within a city, neighborhoods vary in air and water quality levels, residential density and access to parks and open space, exposure to crime and vandalism, and the sociodemographic characteristics of the residents. Within metropolitan areas, there may be several local jurisdictions offering different mixes and levels of

---

[16] See Pope et al. (1979) for an evaluation of various specifications for econometric analysis of land prices.

local public services and different tax rates. Residential choice is not confined to a single metropolitan area: by choosing from among the alternatives offered by the whole country, one may choose levels of all the above-mentioned amenities, usually from a broader range than a single metropolitan area allows. In addition, one may choose climate, topography and vegetation characteristics, and conditions in the labor market.

These choices come at some cost, since mobility is expensive, and it is likely that the opportunity set is more limited than it would be if direct attacks on the nonexclusiveness and nonrivalry problems were successfully executed. Nevertheless, location and land market theory offers useful insights into the linkage between land markets and amenity levels and the opportunities that land markets provide for worker-consumers indirectly to make choices that cannot be made directly.

3.2. Public lands and wilderness areas

The land use theories of Ricardo and Thunen have in common the treatment of land as a factor of production, the assumption that only marketable products matter, and the conclusion that land beyond the fringe (i.e. too infertile in Ricardian and too distant in Thunen theories) will be abandoned to the wilderness. Wilderness land is currently worthless, in these theories, but serves as a reserve of land that may be pressed into service in the event of rising demand for its marketable products.

It is often argued [e.g. Libecap (1981), Baden and Stroup (1982)] that the federal government acquired vast areas of land in the western states during the late nineteenth century under precisely these conditions: it was unwanted and worthless land that would best be placed in the protective custody of government until it was needed for some productive use. This same argument suggest that now – when demands for rangeland, timber, minerals and private-sector recreational developments are impinging on large sections of the public lands – only a stubborn and self-perpetuating bureaucracy prevents the efficient market allocation of these lands to productive uses. The efficiency conclusions of standard land market theory would seem to imply that land publicly withheld from the market is unlikely to be allocated efficiently [Anas (1984)].

While it is reasonable to call for reappraisal of public land policy, Ricardian and Thunen models do not provide an adequate basis for the normative conclusion that land which attracts commercial demands should *ipso facto* be reassigned to the private sector. These models pay no attention to the (often non-exclusive and non-rival) services that such lands can provide when protected from commercial uses that would drastically alter their character. Modern land market theory, as we have seen, recognizes consumer demands for land amenities. Nevertheless, these theories suggest analyses that more readily record the demand for, say, recreational condominiums than for hiking trails or the ecosystem services of wetlands.

Where the services generated by wildlands are predominantly non-exclusive and non-rival, and poorly adapted for value-revelation in land markets, there is good reason to be suspicious of normative uses of land market theory in support of privatization proposals.

3.3. The new communication technologies and choice of residential location

Casual observation has long suggested that the need of most workers to commute to employment centers (the CBD, the factory district, or the transportation hub) limits the choice of residential location. Bid-rent theory focuses on commuting costs and their influence on city form: its compactness or otherwise, and the development of residential zones based on income differences. Changes in transportation technology and pricing policy have clearly influenced city form. It is a commonplace that the increasing availability of automobiles, and the long-standing policy of government at all levels to provide and subsidize super highways, are major causes of the rapid encroachment of suburban and exurban residential developments during the last three decades.

Increasing incomes and relatively lower transportation costs have expanded second-home and vacation-home ownership, with obvious implications for congestion and other pressures on attractive but fragile environments. Some fortunate professionals whose clientele is national – writers and consultants, for example – have long been able to choose where they live so long as telephone service is good and there is a well-served airport within a reasonably short drive. Many of these have chosen year-round homes in the kinds of places where others maintain second or vacation homes.

The new and rapidly developing information technologies seem likely to permit new and substantial categories of workers to complete many of their tasks at home or at small offices far removed from corporate headquarters. Daily commuting will be essential for fewer workers and, for the others, the substantial time and money costs of commuting will be replaced by much lower costs of information transfer. The relative costs of distance from the CBD will be reduced, allowing workers to consume more land and wildlands amenities.

If this scenario comes to pass, as seems likely, markets in land will be less sensitive to distance and more sensitive to amenities, including rural and wildlands amenities. Residential demands will place increasing pressures on farmland and aesthetically attractive coastal, mountain and wildlands areas.

3.4. *Land prices, value capture and the benefits of public programs*

The amenity levels associated with particular parcels of land are reflected in their relative prices. This fact opens the door for hedonic price analyses to estimate the benefits of various environmental improvements and publicly provided services [Strotz (1968), Rosen (1974), and Freeman (1979)]. It also has implications for public finance and equity, briefly considered immediately below.

Suppose that government provides some service that is valued highly by residents and/or businesses near the locational point of provisions but less valued by those more distant. For example, a new subway may be highly valued by residents and businesses within easy walking distance of its stations.

Land market theory predicts that, when a new subway is provided in a previously unserved area, the net gains to nearby residents and businesses will be reflected in increased rents and higher market prices for locationally favored land. If the subway is subsidized by the central government, the number of users enjoying net gains is increased while perhaps a much larger population that enjoys little or no gain is taxed to provide the subway. These circumstances lead many to question the efficiency and equity of providing subsidized, location-specific services financed from general revenues. A commonly proposed solution is to fund such services entirely from user charges.

An alternative, which has the additional advantage that some or all of the public capital outlays are recovered immediately, is presented by land market theory. If locationally-favored land owners experience windfall capital gains as a result of a public project, equity would be enhanced and efficiency in no way impeded by taxing some or all of those capital gains to finance the project. Furthermore, this policy, known as value capture, may provide an alternative method of financing essential infrastructure improvements at a time of strong public resistance to general tax increases. The theory of value capture has been developed by Pines and Weiss (1976, 1982), Wheaton (1977) and Starrett (1981). Anas (1982) discusses value capture procedures and estimates their potential contribution to financing rapid transport systems for Chicago.

3.5. *Macroeconomic policy in a society of real estate owners*

One of the bases of American society has been relatively large participation in real estate ownership, from the yeoman farmers of the eighteenth century, through the homesteaders of the 19th century, and the FHA- and VA-financed suburban homeowners of the 20th century. Real estate ownership has long been considered a key to social stability. In the inflationary 1970s, however, real estate ownership, especially if accompanied by indebtedness, permitted many ordinary citizens (workers and farmers) to acquire unaccustomed wealth. As the analyses of Section 2.5 demonstrate, shocks to macroeconomic policy, as well as to financial and taxation institutions, are quickly translated into windfall gains or losses for those who already own land and real estate.

Ordinary citizens have traditionally been affected by macroeconomic policy via its impacts on employment, wage rates and consumer prices. The emergence of a broad class of citizens with considerable wealth in the form of real estate assets has increased and modified the exposure of citizens to macroeconomic policy. Interest rates are now of concern to many; and the personal impacts of inflation are more ambiguous than ever, due to the relationships between inflation and the asset value of land and real property. Increases (decreases) in expected inflation after purchase are immediately reflected in the increased (decreased) value of the property owner's equity. Macroeconomic policy-makers now have a broader range of concerns when evaluating the effect of policy on the ordinary citizens.

If one takes the endogenous government perspective, one would expect citizens to attempt to influence the conduct of macroeconomic policy in directions that benefit themselves. The emergence of ordinary citizens as holders of significant wealth in real property would, viewed from this perspective, lead one to predict some changes in the macroeconomic policies preferred and promoted by the general public.

3.6. Land and the development of economics

In the introductory section, we noted the acute attention paid to land by the classical economists and the centrality of land in the development of such durable concepts as the distinction between fixed and variable inputs, diminishing productivity of variable factors, and economic rent accruing to the fixed factor. In Thunen's studies of land, largely independent of the classical mainstream of his time, the foundations of marginal analysis and general equilibrium systems are clearly evident.

Considering contemporary mainstream economics, it seems that the Thunen perspective of land may have demonstrated greater staying power than the Ricardian perspective. Modern mainstream economics tends to deny the Ricardian vision of land: the unique fixed resource that is the ultimate and unyielding source of scarcity. Thunen's concepts of location and the cost of distance have proven more durable, although they may be somewhat vulnerable to future cost-reducing technological innovations in transportation and communications. The costs of distance, nevertheless, will remain an organizing principle for economic activity, despite decreasing levels of distance-related costs, so long as important components of cost continue to increase with distance.

While Thunen's concept of land appears more durable than Ricardo's, one must observe that Ricardo's concept once played a more prominent and central role in economics *per se* than Thunen's ever did. Ricardo's concept seemed, at one time, to be crucial to the concept of scarcity and the gloomy predictions of that time about prospects for its alleviation. Thunen's concept was always addressed to a specific question within economics: the location of economic activity.

Both concepts – land as natural resources, and as distance – continue to be the focus of an expanding array of specialized theoretical and empirical endeavors among economists. Advances achieved in the course of this work are quickly integrated into the core analytical technology of the economics discipline. It is in this way that we expect future studies of land economics to make their contribution to the mother discipline.

References

Alonso, William (1964) *Location and Land Use: Toward a General Theory of Land Rent* (Harvard University Press, Cambridge).

Anas, Alex (1982) *Residential Location Markets and Urban Transportation* (Academic Press, New York).

Anas, Alex (1984) "Land Market Theory and Methodology and the Benefits of Public Strategies for Wild Land Management and Allocation", in: G.L. Peterson and A. Randall (eds.), *The Valuation of Wildlands Benefits* (Westview Press, Boulder).

Arnott, R.J. and F.D. Lewis (1979) "The Transition of Land to Urban Use", *Journal of Political Economy* 87, no. 1, 161–170.

Baden, John and Richard L. Stroup (eds.) (1981) *Bureaucracy and Environment* (The University of Michigan Press, Ann Arbor).

Beckmann, Martin J. (1968) *Location Theory* (Random House, Inc., New York).

Beckmann, Martin J. (1969) "On the Distribution of Urban Rent and Residential Density", *Journal of Economic Theory* 1, 60–67.

Beckmann, Martin J. (1972) "Von Thünen Revisited: A Neoclassical Land Use Model", *Swedish Journal of Economics* 74, no. 1, 1–7.

Boyne, David H. (1964) *Changes in Real Wealth Position of Farm Operators, 1940–1960*, Michigan State University Agricultural Experiment Station Technical Bulletin, 294.

Brown, G. and H. Pollakowski (1977) "Economic Valuation of Shoreline", *Review of Economics and Statistics* 69, 273–278.

Castle, Emery N. (1982) "Agriculture and Natural Resource Adequacy", *American Journal of Agricultural Economics* 64, no. 5, 811–820.

Castle, Emery N. and Irving Hoch (1982) "Farm Real Estate Price Components, 1920-78", *American Journal of Agricultural Economics* 64, no. 1, 8–18.

Chicoine, D.L. (1981) "Farmland Values at the Urban Fringe: An Analysis of Sales Prices", *Land Economics* 57, no. 3, 353–362.

Crowley, William D. (1974) "Actual versus Apparent Rates of Return on Farmland Investment", *Agricultural Finance Review* 35, 52–57.

Dempsey, Bernard W. (1960) *The Frontier Wage: With the Text of the Second Part of the Isolated State* (Loyola University Press, Chicago).

Diamond, D.B. and G.S. Tolley (1982) "The Economic Roles of Urban Amenities", in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities* (Academic Press, New York) pp. 3–51.

Dixit, A. (1973) "The Optimum Factory Town", *The Bell Journal of Economics and Management Science* 4, no. 2, 637–654.

Dunn, Edgar Streeter, Jr. (1954) *The Location of Agricultural Production* (University of Florida Press, Gainesville).

Ely, Richard T. and George S. Wehrwein (1964) *Land Economics* (The University of Wisconsin Press, Madison).

Faustmann, M. (1849) "On the Determination of the Value Which Forest Land and Immature Stands Possess for Forestry". English edition in: M. Gane (ed.), *Oxford Institute Paper 42*, 1968, entitled "Martin Faustmann and the Evolution of Discounted Cash Flow".

Ferejohn, John and Talbot Page (1978) "On the Foundations of Intertemporal Choice", *American Journal of Agricultural Economics* 60, no. 2, 269–275.

Found, William C. (1971) *A Theoretical Approach to Rural Land-Use Patterns* (Edward Arnold, Ltd., London).

Freeman, A.M. (1979) *The Benefits of Environmental Improvement: Theory and Practice* (The Johns Hopkins University Press, Baltimore).

Frisch, R. (1931) "Einige Punkte einer Preistheorie mit Boden und Arbeit als Produktionsfaktoren", *Zeitschrift für Nationalokonomie* 3, 62–104.

Gorman, W.M. (1976) "Tricks with Utility Functions". in: M.J. Artis and A.R. Nobay (eds.), *Essays in Economic Analysis. The Proceedings of the Association of University Teachers of Economics. Sheffield, 1975* (Cambridge University Press, Cambridge) pp. 211–243.

Graves, Philip E. and Joanna Regulska (1982) "Amenities and Migration over the Life-Cycle". in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities* (Academic Press, New York) pp. 211–222.

Greenhut, Melvin L. (1970) *A Theory of the Firm in Economic Space* (Meredith Corporation, New York).

Grieson, Ronald E. and James R. White (1981) "The Effects of Zoning on Structure and Land Markets". *Journal of Urban Economics* 10, 271–285.

Grimes, Orville F., Jr. (1982) "The Influence of Urban Centers on Recreational Land Use". in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities*. (Academic Press, New York) pp. 143–164.

Hall, Peter (ed.) (1966) *Von Thünen's Isolated State* (Pergamon Press, London).

Harris, Duane G. (1979) "Land Prices, Inflation, and Farm Income: Discussion". *American Journal of Agricultural Economics* 61, no. 5, 1105–1106.

Harrison, D. and D. Rubinfeld (1978) "Hedonic Housing Prices and the Demand for Clean Air", *Journal of Environmental Economics and Management* 5, 81–102.

Harvey, David (1973) *Social Justice and the City* (Chaucer Press, Bungay).

Helpman, Elhanan and David Pines (1977) "Land and Zoning in an Urban Economy: Further Results", *American Economic Review* 67, 982–986.

Henderson, J. Vernon (1977) *Economic Theory and the Cities* (Academic Press, New York).

Hurwicz, L. and H. Uzawa (1971) "On the Integrability of Demand Functions", in: J.S. Chipman, et. al. (eds.), *Preferences, Utility and Demand* (Harcourt Brace Johanovich, New York) pp. 114–148.

Isard, Walter (1956) *Location and Space-Economy: A General Theory Relating to Industrial Location, Market Areas, Land Use, Trade and Urban Structure* (The MIT Press, Cambridge).

Koopmans, T.C. and M.J. Beckmann (1957) "Assignment Problems in the Location of Economic Activities", *Econometrica* 25, 53–76.

Lee, Warren F. and Norman Rask (1976) "Inflation and Crop Profitability: How Much Can Farmers Pay for Land?" *American Journal of Agricultural Economics* 58, 984–990.

Libecap, Gary D. (1981) *Locking up the Range: Federal Land Controls and Grazing* (Ballinger Publishing Co., Cambridge, for the Pacific Institute for Public Policy Research, San Francisco).

Lösch, August (1954) *The Economics of Location* (Yale University Press, New Haven).

Maser, Steven M., William H. Riker and Richard N. Rosett (1977) "The Effects of Zoning and Externalities on the Price of Land: An Empirical Analysis of Monroe County, New York", *Journal of Law and Economics* pp. 111–132.

Melichar, Emanuel (1979) "Capital Gains versus Current Income in the Farming Sector", *American Journal of Agricultural Economics* 61, no. 5, 1085–1092.

Mills, Edwin S. (1967) "Transportation and Patterns of Urban Development: An Aggregative Model of Resource Allocation in a Metropolitan Area", *American Economic Review* 57, no. 2, 197–210.

Mills, Edwin S. and David M. DeFerranti (1971) "Market Choices and Optimum City Size", *American Economic Review* 61, no. 2, 340–345.

Mirrless, James A. (1972) "The Optimum Town";, *Swedish Journal of Economics* 74, no. 1, 114–135.

Miyao, Takahiro (1981) *Dynamic Analysis of the Urban Economy* (Academic Press, New York).

Pines, D. and Y. Weiss (1976) "Land Improvement Projects and Land Values", *Journal of Urban Economics* 3, 1–13.

Pines, D. and Y. Weiss (1982) "Land Improvements Projects and Land Values: An Addendum", *Journal of Urban Economics* 11, 199–204.

Polinsky, A. Mitchell and Daniel L. Rubinfeld (1977) "Property Values and the Benefits of Environmental Improvements: Theory and Measurement", in: L. Wingo and A. Evans (eds), *Public Economics and the Quality of Life* (The Johns Hopkins University Press for Resources for the Future and the Centre for Environmental Studies, Baltimore) 154–180.

Polinsky, A. Mitchell and Steven Shavell (1975) "The Air Pollution and Property Value Debate", *Review of Economics and Statistics* 57, no. 1, 100–104.

Polinsky, A. Mitchell and Steven Shavell (1976) "Amenities and Property Values in a Model of an Urban Area", *Journal of Public Economics* 5, 119–129.

Pollard, Robert (1982) "View Amenities, Building Heights, and Housing Supply", in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities* (Academic Press, New York) pp. 105–124.

Pope, Rulon D., et al. (1979) "An Evaluation of Econometric Models of U.S. Farmland Prices", *Western Journal of Agricultural Economics* 4, 107–119.

Riley, John G. (1973) "Gammaville: An Optimal Town", *Journal of Economic Theory* 6, 471–482.

Rosen, Sherwin (1974) "Hedonic Prices and Implicity Markets: Product Differentiation in Pure Competition", *Journal of Public Economics* 82, no. 1, 34–55.

Samuelson, Paul A. (1959a) "A Modern Treatment of the Richardian Economy, I: The Pricing of Goods and of Labor and Land Services", *Quarterly Journal of Economics* 73, no. 1, 1–35. in: J.E. Stiglitz (ed.), *The Collected Scientific Papers of Paul A. Samuelson* (MIT Press, Cambridge) pp. 373–407.

Samuelson, Paul A. (1959b) "A Modern Treatment of the Richardian Economy, II: Capital and Interest Aspects of the Pricing Process", *Quarterly Journal of Economics* 73, no. 2, 217–231, in: J.E. Stiglitz (ed.), *The Collected Scientific Papers of Paul A. Samuelson* (MIT Press, Cambridge) pp. 408–422.

Samuelson, Paul A. (1983) "Thünen at Two Hundred", *Journal of Economic Literature* 21, no. 4, 1468–1488.

Schultz, T.W. (1980) "Nobel Lecture: The Economics of Being Poor", *Journal of Political Economy* 88, 639–659.

Schweizer, Urs (1983) "Efficient Exchange with a Variable Number of Consumers", *Econometrica* 51, no. 3, 575–584.

Scott, John T., Jr. (1983) "Factors Affecting Land Price Decline", *American Journal of Agricultural Economics* 65, no. 4, 796–800.

Siegan, Bernard H. (1970) "Non-zoning in Houston", *Journal of Law and Economics* 13, 71–147.

Simon, Julian L. (1981) *The Ultimate Resource* (Princeton University Press).

Smith, Barton A. (1982) "Racial Composition as a Neighborhood Amenity", in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities* (Academic Press, New York) pp. 165–191.

Solow, Robert (1956) "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics* 70, 65–94.

Solow, Robert (1974) "Intergenerational Equity and Exhaustible Resources", *Review of Economic Studies: Symposium on Exhaustible Resources* pp. 29–45.

Solow, Robert M. and William S. Vickery (1971) "Land Use in a Long Narrow City", *Journal of Economic Theory* 3, 430–477.

Starrett, D.A. (1981) "Land Value Capitalization in Local Public Finance", *Journal of Political Economy* 89, no. 2, 306–327.

Strotz, R.H. (1968) "The Use of Land Value Changes to Measure the Welfare Benefits of Land Improvements", in: J.E. Haring (ed.), *The New Economics of Regulated Industries* (Occidental College, Los Angeles) pp. 174–186.

Takayama, T. and G.G. Judge (1971) *Spatial and Temporal Price and Allocation Models* (North-Holland Publishing Company, Amsterdam).

Thünen, Johann H. von (1930) *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, Third edition. Heinrich Waentig (ed.) (Gustav Fischer, Jena).

Tiebout, C.M. (1956) "A Pure Theory of Local Expenditures", *Journal of Political Economy* 64, 416–424.

Varian, H.R. (1978) *Microeconomic Analysis*. W.W. Norton & Co., New York.

Vaughan, Robert J. and Larry E. Huckins (1982) "The Costs of Urban Expressway Noise", in: D.B. Diamond and G.S. Tolley (eds.), *The Economics of Urban Amenities* (Academic Press, New York) pp. 125–142.

Vickerman, R.W. (1980) *Spatial Economic Behavior: The Microeconomic Foundations of Urban and Transport Economics* (St. Martin's Press, New York).

Ward, B. (1972) *What's Wrong with Economics?* (Basic Books, New York).

Wheaton, William C. (1977) "Residential Decentralization, Land Rents, and the Benefit of Urban Transportation Investment", *American Economic Review* 67, 138–143.

Chapter 14

THE ECONOMICS OF FISHERIES MANAGEMENT

GORDON R. MUNRO

and

ANTHONY D. SCOTT*

University of British Columbia, Vancouver

1. Introduction

Fishery resources, unlike minerals and fossil fuels, are renewable in the sense that they are capable of growth. As a consequence, one has the opportunity to harvest such resources indefinitely on a sustainable yield basis and one has further the opportunity to invest (within limits) in the resources simply by harvesting at less than the sustainable yields. As such, fishery resources are similar to forestry resources. Indeed, many parallels can be drawn between fishery and forestry economics (see also Chapters 2 and 12 of this Handbook).

Unlike forests and most other renewable resources, however, fishery resources are difficult to manage effectively because they are, with few exceptions, common property. Most fish, particularly finfish, are very mobile and are, moreover, not readily observable except upon capture. Consequently, it becomes very difficult, or, more to the point, very costly to assign rights of exclusive use to individuals or small groups [Christy (1982)].

The chief economic consequences of the common property nature of the resource are as follows.

(1) If a fishery resource is commercially valuable and is open to unrestricted exploitation, the resource will certainly be subject to excessive depletion from society's point of view. Since the resource is open to all and owned by none, no fishermen will have an incentive to conserve the resource. A fisherman who refrains from harvesting the resource is likely to find, not that he has helped conserve the resource, but rather that he has simply enhanced the harvest opportunities of his competitors.

---

* Funding for this paper was provided in part by the Donner Canadian Foundation through a grant to the Institute of International Relations' research project on "Canada and International Trade" at the University of British Columbia.

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

(2) If the authorities, i.e. the government, should intervene in the fishery to conserve the resource by imposing seasonal or yearly limits on the total harvest, but do nothing to restrict the number of fishermen and vessels competing for the limited harvests, then excess capacity is almost certain to emerge in the fishery.

The problems created by the common property nature of the fishery resources will prove to be central to much of the discussion to follow on the economics of fishery management. We might note in passing that the common property aspects of fisheries create a link between fishery economics and economics of the environment.[1]

2. The role of biological models of the fishery

Every economic model of the fishery, capable of enhancing our understanding of the problems of fisheries management, has as its foundation a biological model of the fishery. One can hardly talk sensibly about growth, sustainable yield, resource investment/depletion without specifying the underlying biological dynamics of the fishery resource in question.

Consider now a fishery involving the exploitation of a single species that does not interact significantly with other species. We refer to the resource in terms of the biomass, the amount of fish expressed in terms of weight. More specifically we focus on the fishable biomass, that part of the biomass which can be exploited, given current technology.

A fishable biomass will be capable of growth as a consequence of:

(a) Recruitment – entry of new fish to the biomass. Fish reproduce. Some fraction of the progeny survive to the fishable age and and thus "recruit" to the fishery.

(b) Growth of individual fish. Individual fish in the biomass experience an increase in body weight as they mature.

The growth of the biomass will be kept in check by mortality:

(a) natural mortality – fish die as a consequence of age, environmental factors and natural predators; and

(b) fishing mortality, i.e. harvesting.

If fishing mortality is absent, the biomass will approach a natural equilibrium at which growth is just offset by the effects of natural mortality.

Denote the biomass (fishable) by x. The percentage rate of growth of the biomass can then be expressed as follows [see Schaefer and Beverton (1963)]:

$$\frac{\dot{x}}{x} = z(x) + g(x) - M(x) - f(E) + \eta, \tag{2.1}$$

where z, g, M, and f, denote the rates of recruitment, growth of fish within the

---

[1] The environment in its capacity as a waste disposal mechanism is subject to over use because it is common property.

biomass, natural mortality and fishing mortality, respectively. The last term on the right-hand side, η, is a random variable with mean zero. Fishing mortality is a function of "fishing effort", E, where fishing effort is taken to refer to the combined flow of labor and capital services devoted to harvesting. An implicit assumption in much of the discussion to follow is that factor proportions are fixed.

The biomass will tend toward equilibrium when (recalling that η has mean zero):

$$f(E)=z(x)+g(x)-M(x), \tag{2.2}$$

$$f(E)x=[z(x)+g(x)-M(x)]x, \tag{2.2a}$$

i.e. when the rate of fishing mortality is equal to the net natural rate of growth of the biomass, or when the harvest, $f(E)x$ is equal to the net natural growth of the biomass in absolute terms.

Ideally, biologists would like to measure all of the rates of growth and mortality entering into eq. (2.1). This has proven to be extremely difficult. Consequently, biologists have found it necessary to introduce various simplifications. This, in turn, has led them to use two broad general approaches, which are commonly referred to as the Beverton–Holt approach – after R.J.M. Beverton and Sidney Holt – and the "general production" or Schaefer approach – after M.B. Schaefer [Schaefer and Beverton (1963)].

In the Beverton–Holt approach, attempts are made to estimate the individual parameters in the context of a discrete time model. The simplifying assumption is usually made, however, that period by period recruitment to the fishery is constant. The behavior of each set of recruits – a cohort or year class – is then examined.

The success of economists in using the Beverton–Holt model as a foundation for economic models has been very limited. If cohorts could be harvested on an individual basis, the economics of harvesting would be straightforward. One would determine the optimal time to harvest the cohort, subject to constraints on fishing effort, in much the same way one would determine when to harvest a forest of uniform age trees.[2]

Harvesting cohorts on an individual basis calls for fishing gear with what the biologists call knife-edge selectivity. Such knife-edge selectivity is the exception, rather than the rule. Almost invariably fishing will occur on a multi-cohort basis. The economics of multi-cohort fishing becomes very complicated indeed. It has, as yet, not really been possible to produce satisfactory analytic solutions, even with the aid of powerful mathematical tools [Clark (1976)].

For this reason fisheries economists have tended to rely heavily upon "general production" or Schaefer type biological models. A Schaefer type model is a "lumped parameter" model in that no attempt is made to distinguish among the

---

[2] See Clark (1976, ch. 8).

factors determining net biological growth. Consequently, net growth of the biomass can be described by the following very simple differential equation:

$$\dot{x} = F(x) - h(t), \tag{2.3}$$

where $F(x)$ is the net natural rate of growth and $h(t)$ is the harvest rate.

The net natural growth of the biomass is thus simply a function of the biomass itself. Properly speaking, net natural growth is a function both of the biomass and of the aquatic environment. The latter is, however, normally assumed to be constant.

The Schaefer model proper is described by the following two equations:

$$F(x) = rx\left[1 - \frac{x}{K}\right], \tag{2.4}$$

where K, a constant, is the natural equilibrium biomass and where r, a constant, is in the "intrinsic" percentage growth rate,[3] and

$$h(t) = qE^{\alpha}x^{\beta}, \tag{2.5}$$

where q, a constant, is the catchability coefficient and where α and β are constants, $\alpha = \beta = 1$.

Equation (2.5) is the harvest production function, a rather special form of the Cobb–Douglas production function. We comment on this further at a later point.

Return to eq. (2.3). Suppose that $x = x^{\dagger}$ and that $h(t) = F(x^{\dagger})$. Obviously the biomass will remain at $x^{\dagger}$, i.e. $\dot{x} = 0$. We would then talk about harvesting occurring at $x^{\dagger}$ on a sustained or sustainable yield basis. Thus for any biomass level, x, $F(x)$ can be seen to represent the sustainable yield associated with x.

Since $h(t)$ is a function of E, as well as x, one can express sustainable yield as a function of E. From the Schaefer model, eqs. (2.4) and (2.5), one can derive the following equation for sustainable yield:

$$Y = uE - vE^2, \tag{2.6}$$

where Y denotes sustainable yield and where $u = gK$ and $v = q^2K/r$.[4]

---

[3] I.e. $\lim_{x \to 0} F(x)/x = r$.

[4] In harvesting on a sustainable yield basis we have:

$$qEx = rx(1 - x/K),$$

from which we can derive an expression for x:

$$x = K\left(1 - \frac{q}{r}E\right).$$

Now substitute for x in eq. (2.5). We then have:

$$Y = qKE - \frac{q^2}{r}KE^2.$$

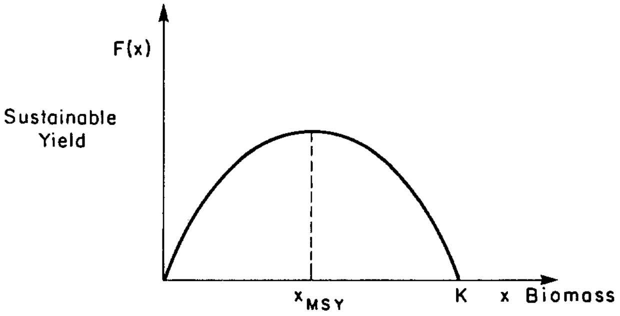

Figure 14.1. Sustainable yield and biomass: from the Schaefer model.

If we graph eqs. (2.4) and (2.6) we get Figures 14.1 and 14.2. The biomass level and effort level associated with the maximum sustainable yield (MSY) we denote by x_{MSY} and E_{MSY}, respectively. It can be easily shown that in the Schaefer model $x_{\mathrm{MSY}} = K/2$.

The management of fishery resources has traditionally been dominated by marine biologists. From about the end of the Second World War the biologists maintained that the appropriate management criterion was that of attempting to achieve MSY or as it was also expressed, "full utilization" of the resource.[5] In terms of the Schaefer model, this would imply permitting the fishery to expand to E_{MSY}, and stabilizing the biomass at x_{MSY}. If the resource were reduced below x_{MSY}, then overfishing (what we shall henceforth call biological overfishing) was deemed to have occurred.

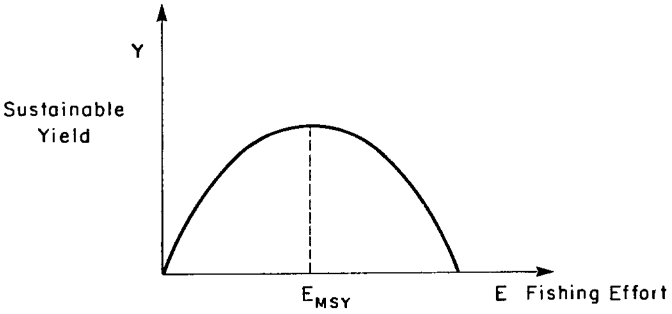

Figure 14.2. Sustainable yield and fishing effort: from the Schaefer model.

---

[5] Larkin (1977).

As we shall see at a later point, economists objected to the MSY criterion because it ignored the costs of harvesting and the true nature of the benefits to be derived from fishing. Partly as a consequence of the campaign waged by economists, fishery resource managers have begun to view the MSY criterion with increasingly less favor.

The Schaefer model has been used extensively by fisheries economists because of its simplicity. Let it be stressed, however, that one can easily employ other general production models. One can, for example, relax without difficulty the restrictive assumption that [see eq. (2.5)] $\alpha = \beta = 1$. One can also, on an *ad hoc* basis, introduce some of the elements of Beverton–Holt type models.

The Schaefer model, as we have presented it, is in continuous time. In much of the discussion to follow we shall make use of continuous time models for ease of exposition. There is no difficulty, however, underlying the use of discrete time, lumped parameter, biological models. Indeed we shall upon occasion turn to discrete time models. Whether one uses continuous time or discrete time models is a matter of taste and convenience [see Clark (1976)].

3. Static approaches to fisheries economics

The origins of modern fisheries economics can be traced back to the seminal mid-1950s article of H. Scott Gordon (1954), "The Economic Theory of a Common Property Resource: The Fishery", which developed an economic model of the fishery using orthodox, static microeconomic analysis. The model was to dominate the theoretical literature for fifteen years and continues to exert an influence on policy-makers up to the present day.

Although not stated explicitly in the article, it is clear that the underlying biological model is a Schaefer type model. Indeed, Schaefer himself provided a firm biological foundation for Gordon's economic model a few years later [Schaefer (1957)]. So close is the link, that the model of the fishery is often referred to as the Gordon–Schaefer model.

The model is one of a single species fishery in which the demand for fish and the supply of fishing effort are both assumed to be perfectly elastic. It is assumed implicitly that the price of landed fish accurately represents the marginal social benefit of harvested fish and that the unit cost of fishing effort is a true measure of the marginal social cost of such effort. There are no troubling second-best conditions to concern us.

With the price of landed fish assumed to be constant, the sustainable yield curves of Figures 14.1 and 14.2 can be transformed into sustainable revenue curves simply by multiplying sustainable yield by the price of fish. Consider Figure 14.3, based upon Figure 14.2. The curve $C(E)$ represents total cost of

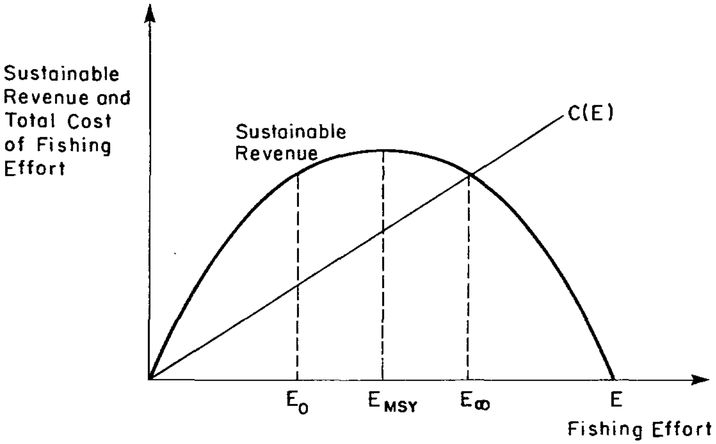

Figure 14.3. Sustainable resource rent maximization and dissipation.

fishing effort where $C(E) = aE$, and where a denotes the unit cost of fishing effort. The difference between sustainable revenue and total cost of effort is defined as sustainable resource rent.

Gordon develops two arguments:

(1) the optimal size fishery is that which maximizes sustainable resource rent, i.e. the fishery should be stabilized at $E = E_0$, and

(2) if the fishery is competitive and is subject to *no* controls, the fishery will expand to the point that resource rent is completely dissipated at $E = E_\infty$

The argument that the optimal size of the fishery is at that point where sustainable rent is maximized rests upon elementary welfare economics. At E_0, the marginal cost of effort, MC_E, is equal to the value of the marginal product of effort, VMP_E. By contrast, at $E = E_\infty$, the $MC_E = VAP_E$, (i.e. the value of the *average* product of effort), or so the argument goes. From this it follows that *economic* overfishing can be deemed to have occurred if E expands beyond E_0.

The second argument that the fishery will expand to E_∞, which Gordon referred to as "bionomic equilibrium", arises from the common property nature of the resource. Suppose that, momentarily, the fishery was at its optimum size, $E = E_0$. Since there is no resource owner (sea lord), the resource rent will accrue to the fishermen and vessel owners, who thus will be earning returns well in excess of their opportunity costs. Given that the fishery is competitive, additional fishermen and vessels will enter the fishery and continue to do so as long as super normal returns are being earned in the fishery. Hence equilibrium will be achieved

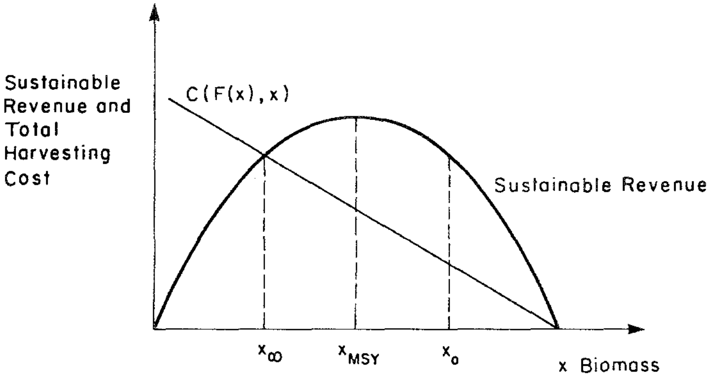

Figure 14.4. Sustainable resource rent maximization and dissipation: a different view.

only at $E = E_\infty$. Consequently, if a common property fishery is subject to no controls, the fishery must invariably expand beyond the socially optimal level.[6]

The case against the policy of using the MSY criterion is also clear. Unless $a = 0$, it will always be true that $E_0 < E_{\mathrm{MSY}}$. Indeed, if effort costs were high enough (or the price of fish low enough) we could have $E_\infty < E_{\mathrm{MSY}}$.

Figure 14.3 is found throughout the literature on the economics of fisheries in which the static analysis is employed. We accompany Figure 14.3 with another figure (Figure 14.4) which is far less common, but which is very useful nonetheless. It is based upon Figure 14.1, with sustainable yield transformed into sustainable revenue. The curve, $C(F(x), x)$ represents the total cost of harvesting the sustainable yield.[7]

The biomass levels x_0 and x_∞ are those associated with sustainable rent maximization and bionomic equilibrium, respectively.

The resource consequences of non-regulation of the fishery are now clear. The resource will be reduced below the optimal level. Indeed, we can now define economic overexploitation of the resource as exploitation which reduces the biomass level below x_0.

---

[6] Except in the trivial case in which the fishery has no commercial potential. Where the fishery does have commercial potential, then the Gordon–Schaefer model predicts, as the reader can verify, that $E_\infty = 2E_0$.

[7] Given that $C(E) = aE$ and $h = qEx$, total harvesting costs are given by $C(h, x) = ah/qx$. If harvesting is taking place on a sustained yield basis, we have:

$$C(F(x), x) = \frac{aF(x)}{qx} = \frac{ar}{q}\left(1 - \frac{x}{K}\right).$$

It will also be true that $x_0 > x_{\text{MSY}}$, that the MSY criterion will lead to economic overexploitation of the resource. This is consonant with our discussion of Figure 14.3, in which we stated that the MSY policy would lead to an overexpansion of the fishery.

In the Schaefer model bionomic equilibrium is achieved because harvesting costs are sensitive to the size of the biomass. The catch per unit of effort (CPUE), i.e. $h/E = qx$, obviously varies directly with the size of the biomass in the Schaefer model. If the resource is depleted, CPUE will fall which implies that unit harvesting costs will rise.[8] Eventually a point is reached that further depletion of the resource would result in negative returns to the fisherman.

There are, however, some fisheries in which the Schaefer model is not applicable and in which CPUE, and hence unit harvesting costs, remain stable as the resource is depleted. This is particularly true of fisheries in which the relevant species has a strong schooling tendency, e.g. herring and anchovies. In such fisheries, complete non-regulation leads, not to bionomic equilibrium, but to extinction of the resource.

The common property problem that we have described, in which complete non-regulation of the fishery results in the dissipation of resource rent, and (from society's point of view) an excessive depletion of the resource, we shall refer to as the Class I common property problem. There are several examples of fishery resource off the Atlantic coast of the United States and Canada that were, it is alleged, subject to an excessive depletion while they held the status of international common property. The depletion of these resources was a major factor in causing both countries to establish unilaterally 200-mile Exclusive Economic Zones.[9]

In the introduction, we referred to a second type of common property problem that arises when the authorities attempt to prevent undue exploitation of the resource by limiting seasonal harvests,[10] but make no attempt to limit the number of fishermen competing for the restricted harvests. The second type of common property problem we hereafter designate as the Class II common property problem.

The Class II common property problem results in dissipation of rent because there will be an excessive number of vessels and fishermen competing for the limited harvests. Fleet redundancy can lead to rent dissipation through "crowding". Vessels may impede one another's movements or disrupt one another's gear,

---

[8] See footnote 7.

[9] At the time of writing, the American zone is referred to as a Fishery Conservation Zone. The FCZ will in time be turned into an EEZ.

[10] The authorities may establish a so-called total allowable catch (TAC) for each such fishery or set "escapement targets", i.e. targets with respect to the number of fish allowed to escape the fishermen and go on to spawn.

thereby leading to economic waste. It is not, however, necessary for crowding externalities to arise for redundancy to lead to rent dissipation.[11]

Too gain further insight into the consequences of fleet redundancy, we turn to a simple, but powerful, model of Clark (1976, 1982) of a seasonal fishery.[12] In the Clark model the following assumptions are introduced:

(1) The authorities impose a seasonal global harvest quota, H. The biomass is stabilized.

(2) The net natural growth of the resource during the fishing season is zero. Harvesting thus reduces the biomass. Growth of the biomass offseason restores the biomass to its original level.

(3) The fishing season has a maximum length, $T_{\max}$. The actual length will depend upon the amount of time required by the fleet to take the allowed harvest (TAC) H.

(4) The fleet, B, consists of uniform vessels carrying uniform gear.

(5) Regardless of the size of B, no "crowding" externalities arise.

Denote the biomass at time t during the season as $x(t)$, $0 \leq t \leq T$. Let $x(0) = X$. The harvest production function is of the usual form: $h(t) = qEx$. During the fishing season, the fleet is in full use. Hence we can replace E with B. Since the net natural growth within season is zero, $h(t) = -\mathrm{d}x/\mathrm{d}t$. Thus we have:

$$\mathrm{d}x/\mathrm{d}t = -qBx, \qquad 0 \leq t \leq T. \tag{3.1}$$

If the allowed harvest, H, is taken over the season, as it certainly will be if the harvest quota has any significance, then

$$H = X(1 - \mathrm{e}^{-qBT}). \tag{3.2}$$

Given B, we shall have

$$T = N/B, \tag{3.3}$$

where $N = (1/q)\log[X/(X - H)]$.

It is obvious that the actual season length, T, will vary inversely with the size of the fleet B. Conversely, since the season length has a maximum length, we can express the minimum fleet size, $B_{\min}$, required to take the TAC, H, over the season as:

$$B_{\min} = \frac{N}{T_{\max}}. \tag{3.4}$$

---

[11] We should note in passing that this aspect of the common property problem is quite different from that experienced in the Class I case. In Figure 14.4, the cost curve represents the *minimum* cost of harvesting the sustainable yield. At x_{∞} in Figure 14.4 the number of vessels and fishermen engaged in harvesting the sustainable yield will be the minimum required. There is no fleet redundancy.

[12] The Class II problem should be analyzed on a season-by-season basis, i.e. discrete time models are called for. Continuous time models are quite inadequate for this purpose.

We can now answer the obvious question, namely: What difference does it make whether B is or is not greater than $B_{\min}$? We can answer this question by considering the *seasonal* cost functions of the individual vessel and of the fleet:

$$c(T) = c_0 + c_1 T, \tag{3.5}$$

$$c(T)B = c_0 B + c_1 TB, \tag{3.6}$$

where c_0 denotes seasonal *non-salvageable* vessels costs and where c_1 denotes *salvageable* costs on a per period of time basis.

The concepts of product specific, or non-malleable, capital, and non-salvageable as opposed to salvageable costs, have gained wide currency in the last few years in the field of industrial organization.[13] Product specific capital implies capital that can be used only in one or a limited number of activities. If these activities cease, the fixed costs associated with the capital cannot be salvaged by shifting to other activities.

In the area of fisheries we can think of fishery specific capital (including human capital). The corresponding seasonal non-salvageable costs, c_0, are to be interpreted as follows. Once a vessel is committed to a fishery at the beginning of the season, c_0 represents those annual costs that the vessel owner cannot "salvage", were the fishery to be suddenly closed, either by ceasing operations or by turning to an alternative activity.

Now return to eq. (3.6) and recall from (3.3) that we have $T = N/B$. Thus, (3.6) may be written as:

$$C(T)B = c_0 B + c_1 N. \tag{3.6a}$$

If all costs were salvageable, then whether the actual fleet, B, did or did not exceed $B_{\min}$ would be a matter of indifference given the absence of crowding externalities. Fleet redundancy would be meaningless and there would be no Class II problem. Thus it is when non-salvageable fleet costs exist, as they almost invariably do, that fleet redundancy takes on meaning.

Let it now be supposed that the authorities have stabilized the resource at a level such that, if harvest costs were minimized, maximum sustainable resource rent would be enjoyed. Let it further be supposed that $c_0 > 0$, but that, momentarily at least, $B = B_{\min}$ and hence $T = T_{\max}$. Let it also be supposed, however, that the authorities place no limits on the size of B.

With $B = B_{\min}$ rent is being generated in the fishery. Since there are no limits on entry, additional vessels and fishermen are attracted into the fishery. Consequently, the authorities, in attempting to prevent the seasonal harvest from exceeding H, steadily shorten the season over time. Vessels lie idle or underutilized during the increasingly long offseason. The fleet will continue to expand,

---

[13] See, for example, Eaton and Lipsey (1981), Klein and Leffler (1981) and Williamson (1983).

however, until

$$pH - c(T)B = 0, \tag{3.7}$$

which can be re-expressed as:

$$pH - \left[c_1 N + c_0 B_{\min} + c_0(B - B_{\min})\right] = 0. \tag{3.7a}$$

The expression $c_0(B - B_{\min})$ we shall refer to as the seasonal *redundancy deadweight loss* in the fishery.

Part of the rent, we should note in passing, may in fact be dissipated through the processing sector, as the problem of non-salvageable costs arises there as well. As the season is reduced in length, the total number of plants required to process the catch will increase. Steady reduction in the season length will mean that plant capital will lie idle or underutilized for increasingly long periods of time.

There are many examples of the Class II common property problem. One such example is provided by the Pacific halibut fishery shared by the United States and Canada. During the first two decades of this century, the resource was subject to increasing and uncontrolled exploitation. The fishery was thus experiencing the Class I common property problem.

Alarmed by the heavy depletion of the resource, American and Canadian authorities joined together in the 1920s to establish what was to become the International Pacific Halibut Commission (IPHC). The IPHC was given a mandate to rebuild and maintain the resource. The IPHC did as it was instructed and by the 1950s appeared to have achieved substantial success [Crutchfield and Zellner (1962)].

The IPHC was given no mandate, however, to control the fleet size. The consequence was that, as the resource was restored, as the effects of the Depression and the Second World War receded, and as resource rent prospects improved, the fleet size grew. The season length decreased *pari passu*.

The halibut fishery being large geographically was divided into areas with harvest quotas being established for each area. In two of the more important areas, Areas 2 and 3, the season lengths declined from 206 days and 268 days respectively in 1933 to 32 and 66 days by 1950. This occurred even though H in the two areas had increased substantially as a consequence of the restoration of the resource [Crutchfield (1982, pp. 38–39)].[14]

The experience of the IPHC raises one other aspect of the Class II problem. Transforming a Class I common property problem into a Class II problem may make society worse off over time. With the Class II, as opposed to Class I, problem there will be management costs that could well be extensive. Once these

---

[14] A full analysis of the economic consequences of the reduced season length is to be found in Crutchfield (1956).

costs are taken into account it may well be the case that when the fleet is at its equilibrium size resource rent will prove to be distinctly negative.[15]

In our discussion of the Class I and Class II common property problems we have suggested that, if either of these problems is allowed to go on unchecked, the inevitable consequence will be that the net economic benefits from the fishery will be fully and completely dissipated and indeed could become negative. This rather depressing conclusion is in fact a product of the linearity of the models used so far and requires some qualification.

If we introduce non-linearities into the models, say by relaxing the assumption that the demand for harvested fish is perfectly elastic or by relaxing the assumption that the supplies of labor and capital are perfectly elastic, then inadequate management will not necessarily result in the complete dissipation of net economic benefits from the fishery. Thus, for example, if the demand for harvested fish is not perfectly elastic, then one can hope that benefits in the form of consumer surplus will be great enough to more than offset management costs [Copes (1972)].

The economist's static model of the fishery has had a decided impact upon North American policy-makers. The impact has come in two stages. In the first stage, fisheries managers, in Canada at least, were prepared to acknowledge the existence of the Class II problem. This led to what one could term a modified MSY policy in which it was thought appropriate to accompany the MSY policy with measures to restrict the fleet size. The first major experiment with limitation on entry to the fishery was in the British Columbia salmon fishery where the Class II common property problem was both stark and obvious [Munro (1982a)].

In the second stage fisheries managers began to abandon the MSY criterion and to replace it with the rather vague concept of Optimum Sustainable Yield (OSY).[16] While vague, the OSY rule did at least allow for economic considerations.

---

[15] If an unchecked Class I common property problem would lead to extinction of the resource, then a case can be made for conserving the resource even if a Class II problem emerges. Before leaving the Class II problem, we should make two comments. The first is that occasionally the problem arises naturally. There are some species the future populations of which are unaffected by current harvesting. Prawns, which normally live for a very short period and which reproduce prolifically, are apparently such a species [Clark and Kirkwood (1979)]. There is no such thing as excessive depletion of the stock so that the Class I problem does not arise. Moreover, there is an upper limit to the seasonal harvest imposed by nature. Lack of government regulation in the fisheries based on such species leads to the redundancy problem.

The second comment is that with fishery specific capital, it is certainly possible that in a completely uncontrolled fishery (Class I problem) the fishery may "overshoot the mark" in the sense that the resource is driven below x_∞. The fleet will continue to operate so long as operating costs are being covered. Obviously this cannot go on forever. The fishery will eventually move to x_∞. Thus with fishery specific capital x_∞ can be seen as the long run equilibrium in a completely unregulated fishery [McKelvey, (1983)].

[16] See, for example, Canada, Department of the Environment, Fisheries and Marine Service (1976); U.S. Fisheries Conservation and Management Act of 1976.

In Canada the application of OSY to that country's large Atlantic fishery resources has clearly resulted in the object of resource management changing from the stabilization of the resources at the MSY levels to levels somewhat greater [Munro (1982a)].

Thus, the static model has achieved some measure of recognition among policy makers. Ironically, as the economist's static model of the fishery was finding increasing favor among policy-makers, it was coming to be viewed with increasingly less favor by the economists themselves.

The reason for the skepticism over the static model's adequacy can be seen by returning to Figures 14.3 and 14.4. In Figure 14.3 it appears that going from bionomic equilibrium to maximum sustainable resource rent is simply a matter of reducing E from E_∞ to E_0, while guarding against the redundancy problem. As Figure 14.4 makes clear, however, that is simply not the case. The resource itself must be rebuilt. Indeed in terms of the underlying model a reduction in E will initially bring about a proportional reduction in the harvest. Achieving future benefits requires that current sacrifices be incurred.

Seen in this fashion, it is obvious that moving from x_∞ to x_0 involves a program of resource investment. It seems equally obvious that the appropriate analysis is capital-theoretic or dynamic as opposed to static. For this reason economists have turned increasingly to dynamic analysis in studying fisheries management problems. In so doing, one might add, they have adopted a mode of analysis that has been used in virtually all other aspects of natural resource economics.

4. Capital theory and the economics of fisheries management

Capital-theoretic analysis came to be applied extensively in fisheries economics by the early 1970s. It must not be thought, however, that economists were unaware of the value of capital theory in fisheries economics before that time. On the contrary, an awareness of the value of capital theory goes back to the beginning of modern fisheries economics. Scott Gordon, whose article provided the foundation of the static analysis of the fishery, gave in 1956 as clear and as forceful a statement on the need for a dynamic approach to fisheries economics as one is likely to find anywhere in the literature.

> The conservation problem is essentially one which requires a dynamic formulation... The economic justification of conservation is the same as that of any capital investment – by postponing utilization we hope to increase the quantity available for use at a future date. In the fishing industry we may allow our fish to grow and to reproduce so that the stock at a future date will be greater than it would be if we attempted to catch as much as possible at the present time...

> In theoretical terms this means that the optimum degree of exploitation of a fishery must be defined as a time function of some sort. That is to say, it is necessary to arrive at an optimum which is a catch per unit of time, and one must reach this objective through consideration of the interaction between the rate of catch, the dynamics of fish populations, and the economic time-preference schedule of the community or the interest rate on invested capital. This is a very complicated problem and I suspect that we will have to look to the mathematical economists for assistance in clarifying it.[17]

A.D. Scott published an article one year after Gordon's 1954 article that represents a pioneering attempt to re-cast the Gordon model in a dynamic framework [Scott (1955)].

The reason that static analysis continued to be used can be found in the Gordon quote. He comments on the difficulty of producing a workable dynamic fisheries model. The task was formidable at the time because existing mathematical tools were inadequate. Attempts were made to construct such models using standard calculus of variations [e.g. Crutchfield and Zellner (1962)]. The resulting models were very complex and difficult to apply. Hence economists tended to retreat to the simpler static models. They may have had misgivings about the static models, but at least they produced results that were comprehensible and thus useful for policy purposes.

The major breakthrough occurred with the development of optimal control theory which can be viewed essentially as an improvement upon and an extension of the standard calculus of variations. Optimal control theory was rapidly applied to modern capital theory (i.e. growth theory) and it was only a matter of time before it came to be applied to fisheries economics as well.

The first serious attempts to apply optimal control theory to fisheries economics appeared in the early 1970s. Pioneering work was done by Plourde (1970, 1971), Quirk and Smith (1970) and others. An extensive and thorough treatment of the subject appeared in 1976 when Colin Clark, an applied mathematician who has developed an extensive interest in economics and biology, published the book: *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*.

While some of the earlier attempts to apply optimal control theory to fisheries economics were daunting in their complexity, the fundamental aspects of dynamic fisheries economics are now sufficiently clear that they are accessible to senior undergraduates [Munro (1981)]. We illustrate with a particularly simple continuous time, deterministic model in which we incorporate all of the assumptions employed in our development of the basic Gordon–Schaefer model.[18] We assume

---

[17] Gordon (1956).

[18] That is to say, we assume that a Schaefer like or "general production" model is the relevant underlying biological model, we assume that both the demand for fish and the supply are fishing effort are perfectly elastic and we abstract from all second-best considerations.

as well that the price of landed fish, the unit cost of effort and the social rate of discount are all independent of time.[19]

The nature of the optimal control problem can be stated as follows. The biomass, $x(t)$, constitutes the state variable, or variable to be controlled. One can control $x(t)$ over time by varying the harvest rate $h(t)$. If the harvest rate is set below the sustainable yield, the biomass will grow, $\mathrm{d}x/\mathrm{d}t > 0$; if the harvest rate is set above sustainable yield, the biomass will decline, $\mathrm{d}x/\mathrm{d}t < 0$. Thus, the harvest rate can be viewed as the control variable.[20] The problem is to control $x(t)$ over time via the harvest rate in such a manner as to maximize the present value of the stream of net economic benefits or returns from the fishery. This will be seen to involve both determining the extent to which society should invest (or disinvest) in the resource and determining the appropriate rate of investment (disinvestment) in the resource over time.

Let us express the flow of net economic benefits, or resource rent, from the fishery at given point in time, t, as

$$\Pi(x,h) = [p - c(x)]h, \tag{4.1}$$

where p and h are as before the price of landed fish and the harvest rate and where $c(x)$ is the unit harvest cost.[21] The objective functional is then:

$$PV = \int_0^\infty \mathrm{e}^{-\delta t}\Pi(x(t),h(t))\,\mathrm{d}t \tag{4.2}$$

where δ is the social rate of discount.

The optimal control problem can now be stated formally as that of determining the optimal control $h(t) = h^*(t)$, $t \geq 0$, subject to eq. (2.3),[22] which we now refer to as the state equation, and to the constraints

$$x(t) \geq 0; \qquad 0 \leq h(t) \leq h_{\max}, \tag{4.3}$$

---

[19] The discussion to follow draws heavily upon Clark and Munro (1975). There are, of course, many alternative treatments of the capital-theoretic approach to fisheries economics. See, for example, Brown (1974), Dasgupta (1982), Dasgupta and Heal (1979), Levhari, Michiner and Mirman (1981), Long (1977), Neher (1974), Peterson and Fisher (1977), and Smith (1977).

[20] Alternatively, one can use fishing effort, $E(t)$, as the control variable. Either is perfectly acceptable. Indeed, at a later point in the chapter fishing effort will be used as the control variable. Generally speaking, the advantage of using the harvest rate as the control variable is that it makes the underlying economics particularly transparent.

[21] As before, we assume that $C(E) = aE$, where a is the unit cost of effort and that the harvest production function is given by $h = qEx$. Thus:

$$C(x,h) = \frac{ah}{qx} \quad \text{and} \quad c(x) = \frac{a}{qx}.$$

[22] I.e. $\dot{x} = F(x) - h(t)$.

where $h_{\max}$ is an arbitrary upper bound [see Clark and Munro (1975)]. Since the objective functional is linear in the control variable, the optimal control problem is itself linear.

We next set up the Hamiltonian of the problem:

$$H = \mathrm{e}^{-\delta t}(p - c(x))h(t) + \lambda(t)(F(x) - h(t)), \tag{4.4}$$

where $\lambda(t)$ is the adjoint or costate variable. The Hamiltonian reflects the benefits from the fishery through current harvests and through benefits to come through investment in the resource. One may interpret the adjoint variable, $\lambda(t)$, as the shadow price of the resource discounted back to the present.

In solving the optimal control problem, one must satisfy the Maximum Principle. One requirement is that

$$\mathrm{d}\lambda/\mathrm{d}t = -\partial H/\partial x, \tag{4.5}$$

which can be re-expressed as

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \mathrm{e}^{-\delta t}c'(x)h(t) - \lambda(t)F'(x). \tag{4.5a}$$

A second requirement is that the Hamiltonian be maximized with respect to the control variable, $h(t)$, at each point in time. Clearly, we shall have

$$\partial H/\partial h = 0$$

only if

$$\mathrm{e}^{-\delta t}[p - c(x)] = \lambda(t). \tag{4.6}$$

Eq. (4.6) is to be interpreted as meaning that we are at a point where the marginal benefit from harvesting is equal to the marginal benefit from investing in the resource. We are then said to be on the so called "singular path".

If (4.6) does not hold, the optimal policy is simply to drive the state variable, $x(t)$, to the singular path at all possible speed. We comment further when we come to discuss optimal approach paths. Prior to investigating optimal approach paths, however, we first investigate the "singular solution" which arises when (4.6) does hold. In so doing we shall determine the optimal level of x and shall derive an investment rule that will prove to be analogous to the Golden Rule of Capital Accumulation familiar from elementary capital theory.

From (4.6) we can say that along the singular path

$$\lambda(t) = \mathrm{e}^{-\delta t}[p - c(x)]$$

and that

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = -\delta\,\mathrm{e}^{-\delta t}[p - c(x)].$$

Hence along the singular path (4.5a) can be written as

$$-\delta\,\mathrm{e}^{-\delta t}[p - c(x)] = -\mathrm{e}^{-\delta t}[p - c(x)]F'(x) + \mathrm{e}^{-\delta t}c'(x)h(t). \tag{4.7}$$

Eq. (4.7) yields the following equation for the singular solution, $x^*(t)$:

$$F'(x^*) - \frac{c'(x^*)F(x^*)}{p - c(x^*)} = \delta \tag{4.8}$$

or, more simply:

$$F'(x^*) + \left.\frac{\partial \Pi/\partial x^*}{\partial \Pi/\partial h}\right|_{h=F(x^*)} = \delta. \tag{4.8a}$$

Eqs. (4.8) and (4.8a) do not involve time explicitly, hence the solution, x^*, is a steady state one,[23] and

$$h^*(t) = F(x^*) \tag{4.9}$$

Equation (4.8), or alternatively (4.8a), is a modified Golden Rule equation and as such provides a rule for determining the extent to which society should invest (or disinvest) in the resource. Note that (4.8) can be re-expressed as

$$\frac{\mathrm{d}\{(p - c(x^*))F(x^*)\}/\mathrm{d}x^*}{p - c(x^*)} = \delta. \tag{4.10}$$

The left-hand side of (4.10) is simply marginal sustainable resource rent resulting from an incremental investment in the resource divided by the cost of the investment, the foregone rent from current harvesting. Thus, the left-hand side of (4.10) is to be interpreted as the yield on the marginal resource investment or the "own rate of interest" of the resource. Thus, (4.10) states that society should invest in the resource up to the point that the own rate of interest of the resource is equal to the social rate of discount.

Equations (4.8) and (4.8a) reveal that the own rate of interest of the resource is divided into two components. These are the instantaneous marginal product of the resource, $F'(x^*)$, and what can be termed the "marginal stock effect". The marginal stock effect $(\partial \Pi/\partial x^*)/(\partial \Pi/\partial h)$ in (4.8a) is a measure of the impact of stock density upon marginal sustainable resource rent. In our model, an increase in the size of the biomass results in a reduction in harvesting costs:

$$C_x(x, h) < 0.$$[24]

If the marginal stock effect were negligible – which is the case in some fisheries – then (4.8) and (4.8a) would reduce to

$$F'(x^*) = \delta. \tag{4.11}$$

---

[23] Is the singular solution, x^*, unique? If the underlying biological model is in fact *the* Schaefer model and if our other assumptions hold, then x^* is unique. Otherwise, there is no assurance that x^* is unique. See Clark and Munro (1975).

[24] See footnote 21.

The similarity of (4.11) to modified Golden Rule equations encountered in elementary capital theory is transparently obvious.[25]

We next consider the question of the optimal approach path to x^*, given that $x(0) \neq x^*$. This really is the question of how rapidly society should invest/disinvest in the resource as it moves toward the target of x^*.

Return to eq. (4.6). If the equation does not hold, the appropriate policy is to set $h = 0$ if $\partial H/\partial h < 0$ or to set $h = h_{\max}$ if $\partial H/\partial h > 0$. The optimal approach path is the so called "bang-bang" approach path.

The implication of the optimal approach path is that, if one is at level of x below x^*, investment in the resource should occur at the maximum rate. Conversely, if one is at a level of x above x^*, disinvestment should occur at a maximum rate. The rational for such an investment policy is that there are no penalties for rapid resource investment (disinvestment).

The absence of penalties for rapid investment arises from the underlying assumptions of the model, both explicit and implicit. The relevant explicit assumptions are the linearity assumptions, in particular, the assumptions that p and a are independent of the harvest rate and the fishing effort rate, respectively. Secondly, there is an important, but implicit, assumption that all of the capital used in exploiting the resource is perfectly malleable. That is to say, none of the capital is fishery specific and hence all costs can be treated as "salvageable".

The assumptions which give rise to the "bang-bang" approach path being optimal are unquestionably restrictive. Clearly we should be prepared to relax them. We shall in fact do just that in the next section.

We are now in a position to make some comparisons between the static and dynamic models. It will be recalled that the static analysis indicated that an optimal fisheries management policy would be one which resulted in maximum sustainable rent, implying that the optimal biomass level would be one at which

$$\mathrm{d}\{[p - c(x)]F(x)\}/\mathrm{d}x\} = 0.$$

It is obvious from (4.10), however, that such a policy would be optimal, $x^* = x_0$, if and only if $\delta = 0$. If $\delta > 0$, which is highly likely, it would simply not be worth society's while to invest in the resource to the extent that resource rent was maximized.[26]

---

[25] What we have termed the marginal stock effect is in fact to be found in standard capital theory where it goes under the name of the "wealth effect". See Kurz (1968). The investment rule we have described arises out of a continuous time model. Clark has shown that there exists a close discrete-time analogue to eq. (4.8) [Clark (1976, ch. 7)].

[26] What about the seemingly plausible rule encountered in the static analysis that one should equate the marginal cost of effort with the value of the marginal product of effort? If one looks more closely at the static model it can be seen that the rule really should be postulated as: marginal cost of effort equal to the value of the marginal *sustainable* product of effort. It is not at all obvious why optimality implies this equality. Indeed, it does not, unless $\delta = 0$.

If it is possible, indeed likely, that x^* will be less than x_0, then what assurance is there that the optimal biomass level will always exceed that associated with MSY, x_{MSY}? In the context of our deterministic model, the answer is none at all. It is quite possible that an MSY policy will prove to be optimal, i.e. $x^* = x_{\mathrm{MSY}}$. It is equally possible that $x^* < x_{\mathrm{MSY}}$. Biological overfishing could be optimal from an economic point of view. Clark (1981) has, for example, argued that, in the case of Antarctic whale populations, social discount rates as low as 1 percent could lead to an $x^* < x_{\mathrm{MSY}}$ outcome.

We should now consider the other extreme, bionomic equilibrium, i.e. $p = c(x_\infty)$. If one returns to eq. (4.10) yet again and multiplies through by $[p - c(x^*)]/\delta$, it can be seen that bionomic equilibrium would be optimal, $x^* = x_\infty$, if and only if $\delta = \infty$.

In any event, since it is virtually inconceivable that the social rate of discount will equal infinity, the dynamic model supports the prediction of the static model that the Class I common property problem will lead to an undue depletion of the resource from society's point of view. We would qualify this support only to the extent of saying that the static model tends to overstate the depletion engendered by an uncontrolled Class I common property problem.[27]

What about the dynamic model's predictions with respect to the Class II common property problem? We postpone discussion of this question until we come to relax the assumptions of perfect malleability of capital.

Dynamic models unquestionably allow one to talk more sensibly about optimal biomass targets. This does not, however, constitute the chief contribution of dynamic analysis to the problems of fisheries management. Rather the chief contribution lies in the fact that the dynamic analysis compels us to recognize the importance of the adjustment phase in fisheries management. As we noted earlier, achieving x^* may require a lengthy and difficult period of adjustment. With the aid of dynamic models the problems associated with the adjustment phase can be analyzed. Static models are, by definition, of no value in studying such problems.

5. Extensions to the basic model

In this section we consider some of the many extensions that have been made to the basic dynamic model. We commence with relatively simple extensions. The model discussed in the previous section was particularly straightforward because it was both linear and autonomous, autonomous in the sense that the parameters were deemed to be independent of time. The first extensions to be considered then involve the relaxation of the restrictive assumptions of linearity and autonomy. We then turn to more complex extensions in which we allow for multiple use of a

---

[27] I.e. $x^* < x_0$.

fishery, for multispecies and finally allow for the existence of non-malleable capital in the fishery.

We commence by relaxing the linearity assumption and shall do so by retaining all of our previous assumptions, except one. It shall now be supposed that the unit cost of effort is aE so that $C(E)=aE^2$ with the consequence that the control model will be non-linear in the control variable $h(t)$. The modified Golden Rule equation will now be:

$$F'(x^*) - \left.\frac{C_x(x^*, h)}{p - C_h(x^*, h)}\right|_{h=F(x^*)} = \delta. \tag{5.1}$$

While eq. (5.1) looks almost indistinguishable from our earlier equilibrium equations, the introduction of non-linearities does have definite consequences. First the optimal approach path is no longer the most rapid. There are now penalties associated with rapid investment or disinvestment. If, for example, $x(0) > x^*$, then a policy of rapid disinvestment would result in very high harvesting costs. Consequently, the approach path will be asymptotic. The appropriate decision rule to apply along the approach path is given by

$$F'(x^*) - \frac{C_x(x^*, h)}{p - C_h(x^*, h)} + \left.\frac{\dot{\Psi}}{\Psi}\right|_{h=F(x^*)} = \delta, \tag{5.2}$$

where $\Psi(t) = e^{-\delta t}\lambda(t)$, which one can interpret as the "current" shadow price of the resources.[28] At $x = x^*$, $\dot{\Psi} = 0$ and eq. (5.2) reduces to (5.1).

The second consequence of non-linearity is that there is no assurance whatsoever that there will exist an unique optimal biomass level. This fact can be seen most easily if it is assumed that $C_x(x^*, h) = 0$, i.e. harvesting costs are independent of the stock level. Eq. (5.1) then reduces to:

$$[p - C'(h)]\ (F'(x^*) - \delta)|_{h=F(x)} = 0, \tag{5.3}$$

and the possibility of multiple equilibria is made transparent.

Next let us restore the linearity assumption, but relax the assumption of autonomy. Let it now be supposed that prices and costs fluctuate through time. We now denote price by $p(t)$ and unit harvesting costs by $\phi(t)c(x)$, where $\phi(t) \geq 0$ is a variable coefficient that allows us to account for shifts in the cost function over time. The equilibrium equation now becomes

$$F'(x^*) + \frac{\partial\Pi/\partial x^*}{\partial\Pi/\partial h} + \left.\frac{\partial\Pi^2/\partial h\,\partial t}{\partial\Pi/\partial h}\right|_{h=F(x^*)} = \delta. \tag{5.4}$$

---

[28] In this instance it is more cumbersome to express the adjoint variable in the form $\lambda(t)$.

In determining the yield on the incremental investment in the resource, one has now to take into account the anticipated immediate change in the marginal return from current harvesting. If the price of landed fish were expected to increase and/or the unit cost of harvesting were expected to fall in the immediate future, this would give one a further incentive to invest now in the resource.

The important thing to note is that the new investment or decision rule is "myopic". That is to say, the decision rule is dependent only upon changes in $\partial \Pi / \partial h$ in the immediate future. It is not necessary to know the full time path of $\partial \Pi / \partial h$. The reason is that investments (disinvestments) made in the resource today are reversible in the future.[29]

The models that we have considered to this point are those of single species fisheries exploited by one group of fishermen. In the real world we find that in some fisheries different groups of rival fishermen are attempting to exploit the resource. Furthermore, it is often found that there are important links between or among species either because fishermen are harvesting several species simultaneously or because there is biological interaction among species. Attempts have been made to extend the dynamic model to deal with both sets of issues.

As an example of the rival exploiters problem, we consider the case of a fishery that is being exploited both by commercial fishermen and by sports fishermen. This is a common occurrence, particularly in the United States. The Washington/Oregon salmon fishery and the California abalone fishery are cases in point.

The difficulty confronting the fishery manager is that there are two distinct sets of net economic benefits being enjoyed from the fishery. The social manager must simultaneously determine the optimal level of investment in the resource in light of these two sets of net benefits and determine how any given level of harvest should be allocated between the rival groups. For assistance in analyzing this problem we turn to a model developed by Bishop and Samples (1980). The authors have both a linear and non-linear version of their model. The non-linear proves to be particularly useful. Hence, it is upon this version that we shall focus.

It is supposed in the Bishop and Samples model that the net economic benefits enjoyed by society through commercial exploitation of the fishery and those enjoyed through recreational exploitation of the fishery can, in some meaningful sense, be compared – no mean feat in practice. Let $h(t)$ denote the harvest rate of the commercial exploiters and $g(t)$ the harvest rate of recreational exploiters. There are now two control variables, i.e. $h(t)$ and $g(t)$.

The social net benefits enjoyed through an incremental commercial and recreational harvest are denoted by $M(x, h)$ and $R(x, h)$, respectively. It is assumed

---

[29] One can, however, run into so-called blocked intervals created by discrete changes in prices or costs that result in the constraints upon the control variable becoming binding. See Clark and Munro (1975).

that:

$$\frac{\partial M(x,h)}{\partial x} > 0, \qquad \frac{\partial M(x,h)}{\partial h} < 0,$$
$$\frac{\partial R(x,g)}{\partial x} > 0, \qquad \frac{\partial R(x,g)}{\partial g} < 0. \tag{5.5}$$

Total net benefits are simply

$$\hat{M} = \int_0^h M(x,u)\,\mathrm{d}u \tag{5.6}$$

and

$$\hat{R} = \int_0^g R(x,v)\,\mathrm{d}v, \tag{5.7}$$

where u and v in this instance serve as dummy variables of integration.

The objective functional is

$$PV = \int_0^\infty \mathrm{e}^{-\delta t}[\hat{M} + \hat{R}]\,\mathrm{d}t, \tag{5.8}$$

while the Hamiltonian for the control problem is

$$H = \mathrm{e}^{-\delta t}[\hat{M} + \hat{R}] + \lambda(t)(F(x) - h - g). \tag{5.9}$$

The Maximum Principle requires that the Hamiltonian be maximized at each point in time with respect to each control variable, $h(t)$ and $g(t)$. Thus, we have:

$$\frac{\partial H}{\partial h} = M(x,h) - \lambda(t) = 0,$$
$$\frac{\partial H}{\partial g} = R(x,g) - \lambda(t) = 0, \tag{5.10}$$

and thus

$$M(x,h) = R(x,g). \tag{5.11}$$

Eq. (5.11) gives a very simple rule for allocating any given harvest between the two groups or sectors. The allocation should be such that the marginal social net benefit to be derived from sports caught fish should be equal to that to be derived from commercially caught fish. The resource investment rule is given by

$$F'(x^*) + \left.\frac{\partial\hat{M}/\partial x^* + \partial\hat{R}/\partial x^*}{\partial\hat{M}/\partial h}\right|_{\substack{h+g=F(x^*)\\ \partial\hat{M}/\partial h = \partial\hat{R}/\partial g}} = \delta. \tag{5.12}$$

The impact of each sector in the determination of x^* is through the marginal stock effect. If the marginal stock effect is negligible, then sectoral considerations will have no influence upon x^*. The sectoral problem will be solely one of harvest sharing.

Finally, let it be noted in passing that, since the model is non-linear, the optimal approach path to x^* will be asymptotic.

The rival exploiters problem proved to be relatively straightforward. The same is not however true of the multispecies problem. Such a problem may arise, as we have suggested, due to the fact that the relevant fleet is multi-purpose in that the vessels shift among fisheries during the year and/or harvest several different species at the same time, or because the species themselves interact.

It is the biological interaction among species that is the source of greatest difficulty.[30] There are innumerable and complex ways in which such interactions may occur. A particular species may compete with another species for food, e.g. sardines vs. anchovies, the species may predate upon another species and in turn may serve as prey for yet another species. Thus, for example, in the Northwest Atlantic capelin serve as prey for cod, but the cod in turn serve as prey for seals. To add to the complexity it is believed that capelin feed on cod eggs [May, Beddington, Clark, Holt and Laws (1979)]. Modelling these interactions in the real world proves to be a daunting undertaking. Indeed, more often than not, the undertaking proves to be impossible.

We shall nonetheless attempt to illustrate some of the economic issues involved by considering an admittedly over-simplified predator-prey model developed by Clark (1976, ch. 9). In the model, it is supposed that both predators and prey are commercially valuable. While the two species interact with one another, it is assumed that they are independent of all other species. It is assumed as well that the species are exploited independently, i.e. there are no by-catch problems.

Denote the biomass of the prey species by x and the biomass of the predator species by y. It is assumed that the fisheries can be modelled as follows:[31]

$$F(x, y) = rx\left[1 - \frac{x}{K}\right] - \alpha xy,$$

$$G(x, y) = sy\left[1 - \frac{y}{L}\right] + \beta xy, \tag{5.13}$$

where r, s, K, L, α and β are constants;

$$\dot{x} = F(x, y) - h_1(t),$$

$$\dot{y} = G(x, y) - h_2(t), \tag{5.14}$$

where $h_1(t)$ and $h_2(t)$ are the rates of harvest of the prey species and predator species respectively.

---

[30] Problems created by multi-purpose vessels and by-catches are difficult, but appear to be somewhat more tractable. See Huppert (1979).

[31] The model is based upon Gause's (1935) model of interspecies competition.

We next assume, for the sake of simplicity, that the demand functions for the two classes of fish are independent and perfectly elastic and that the unit costs of harvesting the prey species are a decreasing function of the size of the biomass, but are independent of the harvest rate. What holds true for units costs of harvesting the prey species, holds true as well for the unit costs of harvesting the predator species. The joint rent function can thus be expressed as:

$$\Pi_{\mathrm{J}}(t) = [p_1 - c_1(x)]h_1(t) + [p_2 - c_2(y)]h_2(t), \tag{5.15}$$

where the subscripts 1 and 2 denote the x and y fisheries, respectively. The objective functional is simply:

$$PV = \int_0^{\infty} \mathrm{e}^{-\delta t}\Pi_{\mathrm{J}}(t)\,\mathrm{d}t. \tag{5.16}$$

It will come as no surprise to learn that there are two joint equilibrium equations [Clark (1976, 318)].[32] They are:

$$F_x + \left.\frac{(\partial\Pi_{\mathrm{J}}/\partial h_2)G_x + \partial\Pi_{\mathrm{J}}/\partial x}{\partial\Pi_{\mathrm{J}}/\partial h_1}\right|_{\substack{h_1=F(x,y)\\ h_2=G(x,y)}} = \delta,$$

$$G_y + \left.\frac{(\partial\Pi_{\mathrm{J}}/\partial h_1)F_y + \partial\Pi_{\mathrm{J}}/\partial y}{\partial\Pi_{\mathrm{J}}/\partial h_2}\right|_{\substack{h_1=F(x,y)\\ h_2=G(x,y)}} = \delta. \tag{5.17}$$

When considering single species fisheries, one can think of the social manager as managing a resource "asset". In the case of a multispecies fisheries, one might think of the social manager as managing a resource asset portfolio. In our example, optimal portfolio allocation requires [from (5.17)] that the marginal yields on the resource investments or own rates of interest of x and y be equal; optimal "investment" in the two resource assets combined requires that the aforementioned marginal yields be equal to the social rate of discount.

The two resource assets are, by definition, interdependent, hence each own rate of interest has an interactive component, $[(\partial H_{\mathrm{J}}/\partial h_2)G_x]/(\partial\Pi_{\mathrm{J}}/\partial h_1)$ and $[(\partial\Pi_{\mathrm{J}}/\partial h_1)F_y]/(\partial\Pi_{\mathrm{J}}/\partial h_2)$, respectively. The expression $(\partial\Pi_{\mathrm{J}}/\partial h_2)G_x$ can be interpreted as a measure of the impact of an incremental change in level of x upon the rent flows from the y fishery. The expression $(\partial\Pi_{\mathrm{J}}/\partial h_1)F_y$ is open to a similar interpretation. Since x is the prey species and y the predator, $G_x > 0$. An increase in the prey population will serve to enhance the predator population. Similarly, we have $F_y < 0$.

To illustrate the significance of the interactive terms, let it be supposed that the prey species were relatively low valued, and that the predator species were

---

[32] There are also constraints $x \geq 0$; $y \geq 0$ which could become binding. If either did so, the analysis would become significantly more difficult. It will be assumed, therefore, that neither constraint becomes binding.

relatively high valued. Antarctic krill and baleen whales provide an example [May et al. (1979)]. The manager would then be given a powerful incentive to "invest" heavily in the prey resource. Let x_0 denote the prey biomass level that would maximize sustainable resource rent if the prey were harvested in isolation, i.e. without reference to the impact of x upon y. When the species interaction is taken into account, it might well prove to be the case that the optimal level of x exceeds x_0.

Conversely, if the predator species were relatively low valued and the prey relatively high valued, e.g. dogfish and Pacific coast salmon, then there would be a strong disincentive to invest in the predator resource. It is quite possible, as Clark points out, that bionomic equilibrium in the predator species fishery would constitute underexploitation of the resource. This can be seen if we rewrite the second of the two equilibrium equations as:

$$\frac{1}{\delta}\left[\left(\frac{\partial \Pi_{\mathrm{J}}}{\partial h_2}\right)G_y+\left(\frac{\partial \Pi}{\partial h_1}\right)F_y+\frac{\partial \Pi_{\mathrm{J}}}{\partial y}\right]=\frac{\partial \Pi_{\mathrm{J}}}{\partial h_2}. \tag{5.18}$$

The left-hand side of (5.18) can be interpreted as the marginal "user cost" of the resource. That is to say, it is the present value of the stream of resource rents that would be lost by an incremental reduction in the biomass through current harvesting.

Let y_∞ denote bionomic equilibrium in the predator fishery. At $y=y_\infty$, we have $\partial\Pi_{\mathrm{J}}/\partial h_2=0$, and (5.18) reduces to:

$$\frac{1}{\delta}\left[\left(\frac{\partial \Pi_{\mathrm{J}}}{\partial h_1}\right)F_{y_\infty}+\frac{\partial \Pi_{\mathrm{J}}}{\partial y_\infty}\right]=0. \tag{5.18a}$$

Since $(\partial\Pi_{\mathrm{J}}/\partial h_1)F_{y_\infty}<0$, it is obvious that marginal user cost could be negative at $y=y_\infty$. If this were the case, then it would be optimal to reduce y below y_∞. Fishermen could not be expected to harvest the resource at a loss, of course. The appropriate policy would be to offer them a bounty to harvest predators.[33]

The determination of the optimal approach paths to the biomass levels, x^*, y^*, is very difficult – and often impossible – to determine [Clark (1976, pp. 319–323)]. Indeed, Clark advocates that one not worry unduly about the optimal approach paths, but rather settle for ones that are "practical" [Clark (1976, p. 323)].

We come finally to relax the assumption of perfect malleability of man-made capital in the fishery. The issue of non-malleable, or fishery specific, capital has already been encountered in our earlier discussion of the Class II common property problem. Indeed the existence of such non-malleable, or fishery specific, capital proved central to the problem.

Now that the malleability assumption is being relaxed, we are in a position to comment on how our perception of the Class II common property problem alters,

---

[33] It should be obvious to the reader that this does not imply that it would necessarily be optimal to drive the predator resource to extinction.

if at all, when using dynamic as opposed to static analysis. The basic answer is that our perception changes very little. We continue to accept the conclusion of the static analysis that, if the Class II problem is allowed to go unchecked, the whole purpose of resource conservation may well be called into question. As the reader can verify, an uncontrolled Class II problem can easily result in the own rate of interest of the resource always tending toward zero.

What we are now in a position to point out, however, is that the existence of non-malleable (fishery specific) capital not only raises the threat of the Class II problem, but has implications for the management of the resource itself. It is to this issue that we now turn.

If the presence of non-malleable capital has important implication for resource management, then we must first ask why it is assumed away in so many dynamic models of the fishery. The answer is straightforward. The assumption of perfect malleability of capital offers immense analytical advantages. Man-made capital can then be treated as a flow-variable with the consequence that there is only one investment problem to be addressed, namely investment in the resource itself. In terms of optimal control theory, we have a relatively tractable one state variable problem. On the other hand, if man-made capital is assumed to be non-malleable, then we are confronted with a much more demanding two state variable problem.

Thus, the cost of relaxing the malleability assumption is high. The returns are also high, however. The implications of relaxing the assumption for the adjustment phase of the resource management program are profound. Indeed, since non-malleable capital is pervasive in fisheries [Baker (1980)], we refuse to relax the malleability assumption at our peril.

We take as an example of the influence of non-malleable capital upon optimal resource management programs the following case. It is assumed that capital invested in the fleet is "quasi-malleable" in that the vessels have no alternative use, have a negligible resale value, but are subject to depreciation. Hence capital can be removed from the fishery gradually over time through depreciation.

It is assumed next that we commence with the fishery in bionomic equilibrium. Finally, it is assumed that when the management program is put into effect, the managers successfully prevent the emergence of a Class II common property problem.

Our model, based on that of Clark, Clarke and Munro (1979), is as follows:

$$\dot{x} = F(x) - qEx, \qquad x(0) = x^0, \tag{5.19}$$

$$0 \leq E \leq E_{\max} = K, \tag{5.20}$$

$$\dot{K} = I - \gamma K, \qquad K(0) = K^0, \tag{5.21}$$

$$I \geq 0, \tag{5.22}$$

$$PV = \int_0^{\infty} \mathrm{e}^{-\delta t}\left[(pqx - c)E - c_K I\right] \mathrm{d}t, \tag{5.23}$$

where K now denotes the fleet size expressed in terms of standardized vessels, $I(t)$ the rate of gross investment in the fleet, $\gamma(t)$ the rate of depreciation, c unit operating costs and c_K, the unit cost of capital. The initial biomass and fleet size are known, being in our example the biomass and fleet size associated with bionomic equilibrium, i.e. $x^0 = x_\infty$; $K^0 = K_\infty$.

The control problem that now confronts us is a two-state variable, $x(t)$, $K(t)$, and two-control variable, $E(t)$, $I(t)$, problem. In this instance it proves convenient to use $E(t)$, rather than $h(t)$, as the control variable relevant to $x(t)$.

The non-malleability of capital assumption is incorporated in (5.22), i.e. there is a non-negativity constraint upon $I(t)$. One can of course physically dispose of vessels. Hence, one can interpret (5.22) to mean that the resale value of vessels, which we shall denote by c_s, is zero. As the reader can easily verify, if c_s were not zero but rather were equal to c_K, our control problem would reduce to a one state variable, one control variable problem.

The fully "synthesized" (or feedback) solution to our control problem is illustrated in the following figure [see Clark, Clarke and Munro (1979)].

Figure 14.5 is a state space diagram (x and K) and is divided by certain curves into three regions R_1, R_2 and R_3. The figure specifies optimal values of I and E for points in these regions. Let us first focus on the biomass levels $\tilde{x}$ and x^* and the curves σ_1 and σ_2.

The biomass levels $\tilde{x}$ and x^* are dual optimal biomass levels that should be viewed as short-run and long-run optimal levels, respectively. Total or "all in" effort costs are given by $C(E) = [c + (\delta + \gamma)c_K]E$. In the long run *all* costs are relevant. When all costs are relevant, then x^* is the optimal biomass level.

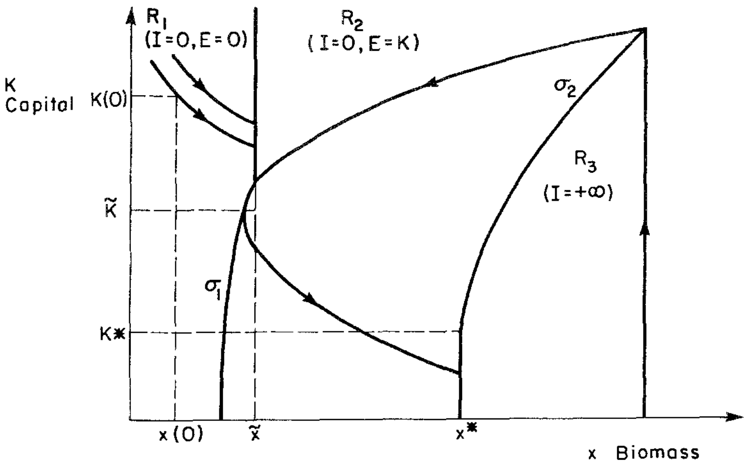

Figure 14.5. Optimal resource and fleet investment when fleet capital is "quasi-malleable".

However, at the beginning of the management program, it will be appropriate for reasons to be given to ignore c_K, i.e. to act as if capital were "free". When capital is "free", and fleet operating costs alone are relevant, the optimal biomass level is $\tilde{x}$. Let us observe in passing that the capital levels K^* and $\tilde{K}$ denote the fleet capacity required to harvest at x^* and $\tilde{x}$, respectively, on a sustained yield basis.

The two curves σ_1 and σ_2 are switching curves. It can be demonstrated [Clark, Clarke and Munro (1979)] that it will never pay to invest in new vessels, i.e. to have positive gross investment in the fleet, when $x < x^*$. Hence gross fleet investment is positive only in R_3. The switching curve σ_2 specifies the appropriate level of gross investment in the fleet. At certain low levels of x, R_1, investment in the resource should proceed at the maximum pace, i.e. $E = 0$. The biomass levels below which E should "switch" to zero is specified by σ_1.

By assumption, the biomass level and fleet size at the commencement of the management problem correspond to bionomic equilibrium. Consider for a moment what the appropriate management policy would be if it were assumed that capital was perfectly malleable. The biomass level x^* would be the sole optimal biomass level. The model is linear in $E(t)$ hence the optimal approach path to x^* would be the "bang-bang" approach path. This implies that the draconian policy of shutting the fishery down entirely until x has grown to x^* should be implemented. Even though investment in the resource would proceed at the maximum rate, $E = 0 = h$, the closure of the fishery could well extend for a long period of time.

Obviously, such a policy could be extremely disruptive to the industry if the capital proved in fact not to be perfectly malleable. What our example demonstrates is that, if capital is not perfectly malleable, the appropriate policy is not draconian, but is rather one of gradual adjustment.

The true optimal management program in our example does call initially for a harvest moratorium, but only until the biomass has grown to $\tilde{x}$, the "free" capital optimal biomass level. The fleet was acquired before the beginning of the management program and has no significant re-sale value. Hence the cost of acquiring the vessels is a bygone and can safely be ignored.

When $\tilde{x}$ is reached $K > \tilde{K}$, capital is "abundant". Optimal policy then calls for harvesting $\tilde{x}$ on a sustained yield basis. The policy is, however, temporary as the "abundance" of the capital is temporary. The depreciation rate, γ, is positive and it is non-optimal to acquire new vessels so long as $x(t) < x^*$. Thus, K must eventually fall below $\tilde{K}$ and $h(t)$ below $F(\tilde{x})$. What one might term an enforced program of resource conservation then comes into effect.

The optimal resource management policy will then be one of positive investment in the resource, but with the rate of investment set at a relatively low level. For it will also be optimal to harvest at the maximum rate with the existing, albeit dwindling, fleet. One intuitive explanation for this policy is that harvesting costs are temporarily low because of the fact that, for a time, only fleet operating costs are relevant. The temporary cost advantages offered thereby are not to be ignored.

Ultimately x^* will be achieved. Then *all* costs become relevant. The fleet capital should be increased to K^* and harvesting should then proceed on a sustained yield basis at x^*.[34]

6. Uncertainty and fisheries management

One of the central characteristics of fisheries management is that it has to be carried out under conditions of great uncertainty. Along with the usual uncertainties associated with price movements and cost shifts through time, there are uncertainties associated with the resource itself. The ability of marine biologists to model fishery resources effectively is, in fact, very limited. Consequently, there will always be environmental "surprises" which will have to be accommodated in fisheries management plans. In applying capital theoretic models to fisheries management problems and thus, by definition, taking time into account explicitly, considerations of uncertainty become inescapable. This in turn has led economists to go beyond deterministic models and experiment with stochastic models.

Two biologists, Walters and Hilborn (1978), list three broad classes of uncertainty in fisheries management. They are concerned largely with environmental uncertainty, but their classification applies as well to uncertainty in strictly economic variables:

(a) random effects whose future frequency of occurrence can be determined from past experience;

(b) parameter uncertainty that can be reduced by research and acquisition of information through future experience; and

(c) ignorance about the appropriate variables to consider and the appropriate form of the model.

Most of the models of fisheries under uncertainty have concentrated on the first and least demanding form of uncertainty. Work is just now underway in dealing with the second class of uncertainty.

In addressing the first form of uncertainty, regardless of the source of uncertainty, one normally starts with a deterministic model. The deterministic model is

---

[34] See also Clarke (forthcoming). Our discussion of extensions to the basic deterministic dynamic model is by no means exhaustive. There are two further extensions of which the reader should at least be aware. First, the discussion to this point has focused on the harvesting sector. Attempts have been made to include the processing sector explicitly in the model. The reader should consult Clark and Munro (1980) and Schworm (1983).

Secondly, in all of the models discussed, we see optimal management leading ultimately to a steady state solution. This outcome is not in fact general. It is quite possible, particularly where economies of scale or startup costs are important in harvesting, that optimal management will call for pulse fishing in which the resource is exploited heavily for a time and then allowed to rebuild. See Hannesson (1975) and Lewis and Schmalensee (1979).

then transformed into a stochastic analogue and the object of management becomes that of maximizing the present value of *expected* returns from the fishery.[35]

Technically, one can set up a stochastic model based on either "open-loop" or "closed-loop" control. Open-loop control implies that all policy actions, present and future, are set at the beginning of the management program. Thus the harvest policy over time would be set at the beginning of the program and would be adhered to come what may. Closed-loop control implies that policy actions are subject to continuous revision as new information becomes available. In practice, it is difficult to believe that any sensible manager would implement the equivalent of open loop controls.[36]

We shall consider two examples of stochastic fisheries models, one developed by Spulber (forthcoming), the other by Charles (1983). Both can be readily linked to deterministic models discussed earlier. Both, it is true, are discrete time models, unlike the continuous time models we have discussed. However this difference is in fact a minor consideration and shall cause us little difficulty.

We turn then to the Spulber model. Uncertainty is introduced through environmental shocks. Prices and costs are assumed to be known through time.

In the absence of environmental shocks we would have the following biological relationship:

$$X_{t+1}=g(X_t), \tag{6.1}$$

where X_{t+1} denotes the resource at the beginning of period $t+1$. Thus, the resource in period $t+1$ is some function of the resource in the previous period. It is simplest to suppose that there is no carry over of the "parent" population from one period to the next. Then X_{t+1} can be viewed unequivocally as the "recruitment" to the fishery at the beginning of $t+1$.[37]

Now introduce harvesting and assume that net natural growth during the fishing season is zero. Net natural growth, if any, occurs offseason. Thus, we have:

$$X_{t+1}=g(X_t-H_t), \tag{6.2}$$

where H_t is the total harvest over the season and (X_t-H_t) is the end of season resource stock or "escapement".[38]

---

[35] For an extensive and thorough survey of work in uncertainty in fisheries management see Anderson and Sutinen (forthcoming).

[36] Note that in a deterministic world the distinction between open and closed loop controls is virtually meaningless.

[37] See Clark (1976, ch. 7).

[38] Let x be the resource stock at a given moment of time within the season, and let the harvest production function be $h=qEx$ and let the season length be τ. Then given there is no intra-season net natural growth, we have:

$$H_t=X_t(1-e^{-qE\tau}).$$

The model is now made stochastic by introducing a disturbance term w_t, where w_t is to be taken as a sequence of independent and identically distributed disturbances. We now have:

$$X_{t+1} = \int_{-\infty}^{\infty} g(X_t - H_t, w_t)\,\mathrm{d}w, \tag{6.3}$$

where X_{t+1} is now the *expected* recruitment for period $t+1$.

Next we introduce a seasonal rent function $\Pi(X_t, H_t)$. It is assumed that the price of fish is constant, that instantaneous unit costs are independent of the harvest rate, and that there are no non-salvageable effort costs. Thus we have:

$$\Pi(X_t, H_t) = pH_t - C(X_t, H_t). \tag{6.4}$$

The term $C(x_t, H_t)$ is the total cost of taking the seasonal harvest. The term can be expressed as:

$$C(X_t, H_t) = \int_{X_t - H_t}^{X_t} c(x)\,\mathrm{d}x, \tag{6.5}$$

where x is the stock at a given moment of time during the season and where $c(x)$ denotes instantaneous unit harvest costs.[39]

From stochastic dynamic programming, an equation for the current value of the resource, X, can be expressed as follows:

$$V(X) = \max_H \left[pH - C(X, H) + \frac{1}{1+r} \int_{-\infty}^{+\infty} V(g(X - H, w)\,\mathrm{d}w \right], \tag{6.6}$$

where r is the discount rate and $1/(1+r)$ is the discount factor.

Spulber is able to show that the stochastic analogue to the deterministic Golden Rule equation is given by:[40]

$$\int_{-\infty}^{+\infty} g_X(X - H, w)\,\mathrm{d}w + \int_{-\infty}^{+\infty} \frac{[\partial\tilde{\Pi}/\partial H - \partial\Pi/\partial H + \partial\tilde{\Pi}/\partial X]}{\partial\Pi/\partial H} \cdot g_X(X - H, w)\,\mathrm{d}w = 1 + r, \tag{6.7}$$

where $\tilde{\Pi}$ denotes *expected* rent from the fishery. Thus, (6.7) can be interpreted simply as stating that one should invest in the resource up to the point that the expected yield or return on the resource investment is equal to the appropriate interest rate.

---

[39] See footnote 37.

[40] Properly speaking, it is the analogue to the *discrete-time* deterministic Golden Rule equation. See Clark (1976, p. 245).

One can also express (6.7) as:

$$\frac{1}{1+r}\int_{-\infty}^{\infty} g_X(X-H,w)\left[\frac{\partial\tilde{\Pi}}{\partial H}+\frac{\partial\tilde{\Pi}}{\partial X}\right]\mathrm{d}w=\frac{\partial\Pi}{\partial H}. \tag{6.7a}$$

Thus, one should invest in the resource up to the point that the present value of the expected stream of net benefits from incremental investment in the resource, or alternatively expected user cost, is equal to the marginal net benefit from current harvesting.

Once the target has been achieved, the optimal harvest rate will not be identical season after season as the deterministic model would suggest. Rather the optimal harvest rate will vary in response to the environmental shocks to which the system is subject. As Spulber stresses, the model suggests, not a sustained yield equilibrium, but rather a time invariant probability distribution of harvest levels.

One important question that emerges is whether the *expected* optimal biomass level is greater than, less than, or equal to the corresponding deterministic biomass equilibrium. There is no clear-cut answer. Nonetheless, the work of Reed (1979) and Lewis (1981) imply that under reasonable cost and revenue assumptions, the expected optimal biomass level will generally be greater than its deterministic analogue. With environmental uncertainty there is an upside risk that one will find the resource stock is larger than desired, i.e. it will be wished that more intensive harvesting had occurred. There is also a downside risk that the resource stock will prove to be smaller than desired. The consequence may be that a lengthy and painful period of stock restoration will be required. The Reed and Lewis results suggest that the downside risk tends on balance to outweigh the upside risk.[41]

One question that is not addressed in the articles cited to this point, but which is important nonetheless, is how the inescapable resource variability is to be dealt with once the managers are "on target". Biologists point out [e.g. Doubleday (1975)] that one can work towards a stable biomass, a stable season by season harvest or a stable season by season effort rate. There is a clear tradeoff between the size of the average seasonal harvest and variability of harvests. Working toward a stable biomass level will produce the largest average seasonal yield, but at the expense of large seasonal variation in harvests. Working towards a stable seasonal harvest will, by definition, reduce harvest variability, but will do so at the cost of significantly reducing average seasonal harvest. A policy of stable seasonal effort rate will produce a compromise outcome [Doubleday (1975)]. There is no *a priori* answer as to where along the tradeoff frontier it is appropriate to be. The problem is analogous to that of yield versus risk common in financial theory. A similar sort of analysis is required [see Mendelssohn (1979)].

---

[41] Interestingly, Reed (1979) argues that this will be true even if the manager is risk neutral.

We turn next to the Charles (1983) model. Like the Spulber model, it is a seasonal model in which uncertainty is introduced through environmental shocks. However, the Charles model goes beyond the Spulber model (or those of Lewis and Reed) in that it constitutes a stochastic analogue of the Clark, Clarke and Munro (1979) model involving non-malleable capital. As a consequence, Charles must deal with two inter-related investment problems under uncertainty, i.e. investment in the resource and investment in the fleet. He assumes that the capital is quasi-malleable in the sense that we used this term in the previous section. The vessels have a negligible re-sale value, but the fleet can be reduced through depreciation.

Charles finds that, as far as resource investment is concerned, it will by and in the large be optimal to invest more heavily in the resource than it would in a deterministic world. As such, his results are consistent with those of Lewis, Reed, and others. The real question, however, is whether optimal investment in fleet capacity will be greater than or less than it would be in a deterministic setting.

The non-malleability aspect of the fleet capital introduces both a downside and upside risk to investment in fleet capacity. The harvest opportunities available to the fleet at the beginning of each season will depend on the seasonal "recruitment" to the fishery which is, of course, subject to uncertainty. There is the downside risk that, in periods of exceptionally low recruitments, the fleet will prove to be too large in the sense that a large part of the fleet will be idled. There is also an upside risk, however, that, in periods of exceptionally large recruitments, the fleet will prove to be too small to take full advantage of the harvest opportunities.

Whether the optimal fleet under uncertainty will be larger or smaller than the optimal fleet under certainty, will depend in part on the relative importance of non-salvageable, as opposed to salvageable, costs in fleet operations. The more important are non-salvageable costs, the more severe obviously will be consequences of fleet underutilization.

The size of the optimal fleet under uncertainty in relation to that under certainty will depend as well, however, upon the relative rate of growth of the resource. One important characteristic of fast growing resources is that the harvest consequences of exceptionally large or small recruitments are shortlived. This means, on the one hand, that if the harvest opportunities offered by an exceptionally large recruitment are not taken promptly they may be lost forever. Similarly, the harvest consequences of low recruitments can be expected to be transitory. Slow growing resources, on the other hand, have a longer "memory". The effects of large recruitments will be long lasting; harvest opportunities lost today may, in part at least, be there tomorrow. The effects of exceptionally low recruitments will also be long lasting. Thus, with fast growing resources the upside risk of fleet investment will be emphasized. The downside risk will be de-emphasized. With slow growing resources the reverse will be true.

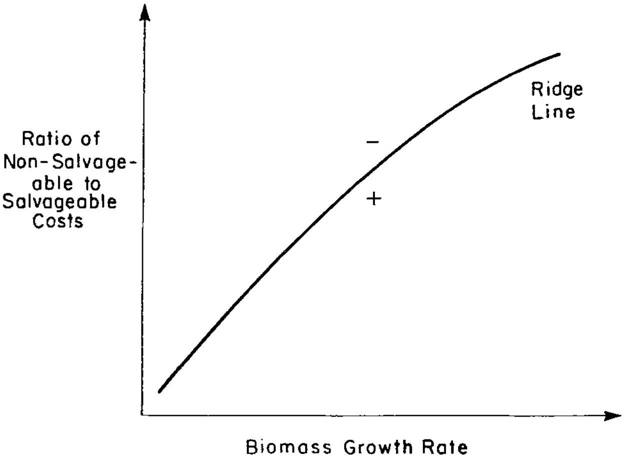

Figure 14.6. Optimal fleet investment and uncertainty: upside versus downside risks.

The relationship between rate of growth of the resource and the relative importance of non-salvageable costs can be illustrated with the aid of Figure 14.6 adapted from Charles (1983). The ridge line is the locus of all combinations of cost ratios and biomass growth rates at which the upside and downside risks are evenly balanced. Hence, along the ridge line, optimal fleet investment policies under conditions of uncertainty and certainty are identical. At points above the ridge line the downside risk predominates; optimal fleet investment is less under uncertainty than under certainty. Conversely, below the ridge line the upside risk predominates; optimal fleet investment under uncertainty is greater than under certainty.

We conclude by noting that much work remains to be done in the area of uncertainty. The very important problem of what constitutes optimal economic management of the fishery when the underlying biological parameters are simply unknown has yet to be addressed by economists.

7. Major policy issues I: The regulation and control of harvesting capacity

The common property problem, which we have now analyzed at some length, leads to excessive harvesting capacity. In the classic or Class I form of the common property problem, where there is a complete absence of government regulation, excessive harvesting capacity arises in the sense that the fleet becomes sufficiently large to reduce the resource well below the optimal level.[42] Either the resource is driven to extinction and the fishery destroyed, or harvesting costs are raised as a consequence of the resource depletion to the point that net economic benefits from the fishery are more or less fully dissipated. In the Class II form of the problem there is excessive harvesting capacity in the sense that some part of the capacity is genuinely redundant. Redundancy alone can lead to complete dissipation of rent. Economic waste may also arise, however, through crowding externalities.[43]

The basic policy problem then is to control harvesting capacity. In a new fishery, this means devising measures to prevent excessive capacity from emerging. In an old fishery where either the Class I or Class II problem has been allowed to go unchecked in the past, it means devising measures, first to eliminate the present unwanted capacity and secondly to prevent the re-emergence of excessive capacity in the future. It goes without saying, that the old fishery case is far more difficult to deal with than the new fishery case. The political and social problems to be encountered in eliminating unwanted capacity can prove to be formidable.

In our discussion of proposed policy measures for controlling capacity, we shall confine ourselves largely to the Class II common property problem. Some of the policy measures that have been put forward can, in theory, be used to address both Class I and Class II problems simultaneously. It is difficult, however, to believe that anyone would seriously consider demanding so much of these measures. It is far more reasonable to suppose that the authorities will attempt to prevent excessive depletion of the resource through the establishment of total allowable catches or escapement targets (as they do now) and then turn to the other measures we shall discuss, to deal with the fleet redundancy problems.[44] Indeed the most promising measure we shall discuss is clearly one to be applied only in conjunction with TACs or escapement targets.

In a fishery in which the Class II common property problem has emerged, or threatens to, the authorities can attempt to deal with the excess harvesting capacity problem either through the use of input controls or through the use of output controls [Scott and Neher (1981)]. Input controls are, in theory at least,

---

[42] See Figures 14.3 and 14.4.

[43] We have talked as if there is a clear separation between the Class I and II problems. This is somewhat of an overstatement. If there is substantial fleet redundancy, this is certain to make the authorities' resource monitoring task more difficult.

[44] If we commenced with a fishery in bionomic equilibrium, the authorities would introduce measures to rebuild the stock. If the capital in the fleet exhibits any non-malleability, fleet redundancy would appear as soon as the stock rebuilding measures were introduced. From our earlier analysis, we know that the appropriate policy would be not one of wholesale reduction of the fleet, but rather one of allowing the fleet to decrease over time through depreciation. As conditions in the fishery improve, the authorities would have to be on guard against the emergence of the Class II problem.

very simple. In a fishery where the redundancy problem has yet to become serious the authorities might establish a program of limited entry in which they restrict the number of vessels (and fishermen). The authorities could, for example, stipulate that no vessels will be permitted to operate without a license and then proceed to limit severely the number of licenses. Once the fleet is in place, the authorities could then further stipulate that a new vessel would be allowed to join the fleet only if an existing member of the fleet were simultaneously removed.

If fleet redundancy were already severe, the authorities would first have to eliminate excess capacity. One method that has been attempted in practice is for the authorities to set up a scheme to buy up or "buy-back" vessels in the fleet. The vessels purchased by authorities could then be sold to non-participants in the fishery. A limited entry program would accompany the buy-back to prevent the re-emergence of excess harvesting capacity.

For reference purposes, let us recall eq. (3.7a) designed to illustrate full rent dissipation in the Class II problem case:

$$pH - \left[c_1 N + c_0 B_{\min} + c_0(B - B_{\min})\right] = 0$$

where $c_0(B - B_{\min})$ is the redundancy deadweight loss. Suppose that in the fishery in question, eq. (3.7a) held. The purpose of the buy-back would be to reduce the actual fleet size B to $B_{\min}$.[45] Entry controls would then be used to prevent B from becoming larger than $B_{\min}$.

One assumption underlying (3.7a) is that all vessels (and fishermen) are identical. If this were indeed the case, then the seasonal rent enjoyed by the fleet and the individual vessel owner could be expressed as follows:

$$pH - [c_1 N + c_0 B_{\min}] = \Pi_{\mathrm{SL}} \tag{7.1}$$

and

$$p(H/B_{\min}) - C(T_{\max}) = \Pi_{\mathrm{SL}}/B_{\min}, \tag{7.2}$$

where Π_{SL} is the seasonal fleet resource rent. The individual vessel owner/fisherman would still have every incentive to maximize his share of Π_{SL}. If, however, it was true that all vessels were identical, that input substitution was impossible, further that technology was frozen over time, the vessel owner/fisherman would be thwarted in his desire. He might succeed in buying a vessel license from another vessel owner. This would avail him little, however, as the license price would certainly reflect the capitalized value of expected rents over time.

---

[45] Return to Figure 14.5. Let it be supposed that the authorities have stabilized the resource at x^*, but that the fleet capacity is $\bar{K}$. Thus there is capacity redundancy equal to $\bar{K} - K^*$. Suppose, further, that the authorities are now in a position to exercise iron control, not only over the resource, but over the fleet as well. What would be the correct policy? It would *not* be to buy-back the excess capacity. It would rather be to "unleash" the fleet and allow the resource to be depleted. The resource would gradually rebuild as fleet capacity declined through depreciation. Fleet reduction through attrition, rather than through a buy-back program, would be appropriate.

If, however, vessels are not alike and, more importantly, if input substitution is feasible and if technology is not frozen over time, the situation changes entirely. Then vessel owners would be able to respond by seeking to upgrade the catching power of their existing vessels or of "trading up" in terms of vessels. This could not be done costlessly, of course, but so long as the vessel owner's extra capital cost was less than the present value of the stream of additional rent he could expect to enjoy, it would pay him to upgrade or trade up. Thus, unless the authorities were able to police vessel owners effectively to ensure that they did not circumvent the regulations, harvesting capacity would be certain to creep upwards inexorably even though the *number* of vessels remained constant or even declined. The inexorable upward increase in harvesting capacity could well result in the buy-back/limited entry program providing no more than a temporary abatement of the Class II problem.

With fleets consisting of a few large high profile company owned vessels, as is typical of offshore groundfish fleets in the Canadian Northwest Atlantic, policing could well prove to be adequate.[46] With fleets consisting of hundreds and perhaps thousands of small individually owned vessels, which is typical of inshore and nearshore fisheries on both Atlantic and Pacific coasts of North America, effective policing of vessel owners proves to be very difficult, if not impossible.

The British Columbia salmon fishery, which is a small boat fishery, provides an interesting example of the difficulties to be encountered in limited entry programs. In the late 1960's, the Canadian authorities introduced a limited entry program to the aforementioned salmon fishery, which was at the time suffering from an acute case of the Class II problem [Anderson (1977)]. A strict vessel licensing system was introduced along with a buy-back program. By 1980 the limited entry program had apparently achieved some considerable measure of success. The total number of vessels had declined by 20 percent [Canada, Commission on Pacific Fisheries Policy (1982)]. The success was illusory, however. It is estimated that by the late 1970s the amount of capital employed in the industry may in fact have increased by as much as 50 percent [Fraser (1979)] over the previous decade.

When the limited entry program was originally introduced, a boat-for-boat rule was put into effect. Anyone wishing to bring a new vessel into the fleet had first to acquire the license of an existing vessel in the fleet. Once the existing vessel was stripped of its license, it was barred from the fleet.

Fishermen responded by replacing small vessels with larger ones, which had, of course, correspondingly greater harvesting power. The authorities responded by changing the boat-for-boat rule to a ton-for-ton rule. This quickly led to "pyramiding" in which sets of small vessels were removed from the fleet and their licenses combined to bring in larger vessels with greater catching power per ton.

---

[46] See Mitchell, (1978).

After several years the authorities introduced new rules to curb pyramiding. There has not yet been time to assess the effectiveness of the new rules. There are no obvious reasons, however, why one should be optimistic.

In assessing the entire Canadian experience with input controls, Scott and Neher (1981, pp. 31–32) conclude that "fishermen have generally been more adept at evading the original intent of input restrictions than the regulators have been at devising and imposing them". There are, the authors conclude, too many inputs for all to be controlled effectively. Hence opportunities for substitution are never entirely absent [Scott and Neher (1981)].

The common property problem in either of its forms is essentially one of market failure. The market sends out incorrect signals to the participants in the fishery. Input controls constitute an attempt to address the problem by making it difficult for participants to respond to the incorrect market signals. Output controls, on the other hand, change the market signals themselves. With the inadequacies of input controls manifest, economists have been giving greater emphasis to output controls.

Two forms of output controls have been emphasized, landings taxes and individual fisherman (or company) quotas. The rationale for the tax, or royalty, is straightforward. If rent from the fishery can be absorbed by a tax, the tendency toward overexpansion of the fleet will be curbed.

The effect of the tax can be illustrated if we return to eqs. (7.1) and (7.2) and assume, for the moment, that all of the underlying assumptions are valid. We have Π_{SL}, the seasonal rent that would be enjoyed by an untaxed fleet of size $B_{\min}$. Now introduce a per unit landing tax equal to Π_{SL}/H, a tax that will be perceived by the fishermen as a reduction in the price of harvested fish [Strand and Norton (1980)]. The rent perceived by the fishermen will vanish, i.e.

$$(p - T_x)H - [c, N + c_0 B_{\min}] = 0 \tag{7.3}$$

and

$$(p - T_x)H/B_{\min} - C(T_{\max}) = 0, \tag{7.4}$$

where, by assumption, $T_x = \Pi_{\mathrm{SL}}/H$.

Vessel owners now have no incentive to try to increase their harvesting capacity. Indeed, if we now allowed for different classes of vessels, input substitution and technological change through time, the tax could always be set so that only the most efficient would survive.

If initially we had $B > B_{\min}$, the introduction of tax $T_x = \Pi_{\mathrm{SL}}/H$ would obviously result in vessel owners enjoying perceived rents of less than zero. They would go on enjoying negative rents until the fleet had been reduced to $B_{\min}$. Thus, the imposition of the tax would be an effective, albeit brutal, means of eliminating redundant capacity. Once the redundancy had been removed, potential participants would be on the margin of indifference with respect to entry.

Excessive optimism about the future of the fishery could, it is true, lead to excessive entry. The presence of the tax would ensure, however, that such unwarranted optimism would be met with swift punishment.[47]

The aforementioned tax policy for ridding the fishery of redundant capital has one drawback. It is politically impossible. We can, however, do somewhat better than this. One can suggest a policy in which the excess capital in the fishery would be removed through a combined tax and buy-back program. The tax would be set in such a manner that the resource rent accruing to the fishermen would be equal to zero, but would not be negative.

Let

$$\gamma = \frac{c_0(B - B_{\min})}{\Pi_{\mathrm{SL}}},$$

where, as before, Π_{SL} is the seasonal rent from the fishery that would be forthcoming when the fishery is being managed with maximum efficiency. Now set the tax, T_X, such that:

$$T_X = \theta \frac{\Pi_{\mathrm{SL}}}{H}, \tag{7.5}$$

where $\theta = 1 - \gamma$.

Suppose now that when the fleet reduction program commences, eq. (3.7a) holds, i.e. the entire potential rent from the fishery is being dissipated through fleet redundancy. Thus, we shall have $\gamma = 1$ and the appropriate level of T_X will be zero. As the buy-back program is put into effect, however, γ will steadily decline and thus T_X will steadily increase until $B = B_{\min}$, $\gamma = 0$ and $T_X = \Pi_{\mathrm{SL}}/H$. The reader can verify that, throughout the fleet reduction program, the resource rent accruing to the fishermen will be equal to zero, neither more nor less. One might add in passing, that the zero resource rent for fishermen policy will have a favorable impact upon the cost to the authorities of the buy-back program.[48]

While the policy of phasing in the tax would be more palatable than the earlier tax policy we described, there are still difficulties with taxes as a measure of control. The first is that even when applied in a non-brutal manner, one can anticipate strong political opposition to the introduction of taxes. Fishermen are not noted for their reticence in using any and all political power at their command.

---

[47] It can be shown that under certain circumstances a tax, or set of taxes, can deal with both Class I and Class II problems simultaneously, e.g. see Brown (1974) and Clark and Munro (1980). However, this outcome should be regarded for what it is – a theoretical curiosity.

[48] If it were feasible to introduce a program of fleet attrition, rather than fleet buy-back, as described in footnote 45, taxes could be used in the same manner. That is to say, taxes could be phased in, once fleet attrition commenced, so as to keep resource rent accruing to the fishermen equal to zero. The one qualification is that as x^* was approached, K would certainly be less than K^*. Measures would have to be undertaken to ensure that there was a once and for all positive net investment in the fleet sufficient to raise it to K^*.

Secondly, the prices and costs are certain to fluctuate over time, which will make effective administration of the tax difficult. If the tax is set too high, the full allowable harvest will not be taken; if it is set too low, the problem of excess capacity will re-emerge. If the harvests of individual vessels consist of a variable mix of species, the difficulties will be compounded [Scott and Neher (1981)].

The second form of output control consists of establishing a system of harvest quotas for individual fishermen, an approach first discussed in detail by Christy (1973). Under this approach, the authorities determine H, the optimal seasonal harvest as before, and then subdivide H among individual fishermen. The quotas might be distributed on the basis of some formula or alternatively might be auctioned off by the authorities. If the fleet consisted of uniform vessels, was at its optimum size, and if the quotas were distributed equally, then each quota would be simply $H/B_{\min}$.

In establishing a system of fishermen quotas, the authorities would, in effect, be establishing a system of property rights for the fishermen. The property rights would be rights, not to the resource itself, but rather to segments of the harvest.[49] Some authors, e.g. Moloney and Pearse (1979), liken the system of harvest quotas to the medieval practice of "stinting" in which peasants were given rights to allow specific numbers of livestock to graze on the feudal common. Modern examples of stinting can be found in the administration of rangelands, oil reserves, timber lands and of water resources [Moloney and Pearse (1979, p. 860)].

Unlike a global harvest quota, there will be no worry about excess vessel capacity unless, that is, cheating is possible. If cheating is possible through fishermen taking and delivering unrecorded catches, then the system will break down and the common property problem will re-emerge in full force.

In the absence of cheating, however, fishermen will have no incentive to expand harvesting capacity, other than through buying quota rights from fellow fishermen. Moreover, the authorities need have no concern about the types of vessels used. Each fisherman will have an incentive to use the most efficient harvesting means available.

The greatest efficiency will be achieved if the quotas are made transferable, i.e. marketable. Those fishermen that are less efficient than their fellow fishermen – lowliners vs. highliners – will have an incentive to sell their quota rights to their more efficient competitors.

Clark (1980) proves that, if fishermen quotas are freely transferable, the quota system will have the same effect in efficiency terms as taxes. To would be entrants to the fishery, the cost of acquiring a seasonal quota will be comparable to the

---

[49] It has been suggested that, in some circumstances, it might be possible to assign quota rights, not to the harvest, but to the resource itself [see Scott (forthcoming)]. If this were indeed possible, it would go right to the heart of the common property problem in both its Class II *and* Class I forms. The fishermen would become a participant in resource management with a firm stake in the future, as well as current, returns from the resource. Growing interest, particularly in world salmon fisheries, in ocean ranching, fish farming and terminal fisheries, suggests that the possibility is by no means fanciful.

cost arising from an efficiently administered tax. To those fishermen with quotas, there will also be a cost comparable to a tax in the form of an opportunity cost of holding the quotas.

Politically, of course, individual quotas are far more palatable than taxes. In one fishing nation at least, Canada, individual quotas have found considerable favor with non-academic economists and have in certain fisheries actually been implemented [Scott and Neher (1981)]. The year 1982 saw the submission of two major reports on Canada's Atlantic coast and Pacific coast fisheries, both commissioned by the federal government [Canada, Task Force on Atlantic Fisheries, Report (1982); Canada, Commission on Pacific Fisheries Policy, Final Report (1982)]. Both reports, particularly the former, advocated the use of individual fishermen's quotas as an effective means of regulating the fishery.

Both reports, however, also pointed to major limitation on the use of fishermen's quotas. If the seasonal harvests are extremely variable, and unpredictable, the administrative costs and difficulties of running such a system become severe. Thus, while the authors of the aforementioned report on Canadian Pacific fisheries were strong advocates of fishermen's quotas, they refused to recommend a quota system for the predominant salmon fishery because of the wide unpredictable fluctuations to which harvests in this fishery are subject [Canada, Commission on Pacific Fisheries Policy, Final Report (1982, ch. 9)].

Our understanding of the effects of various regulations upon fisheries is far from complete. One can speculate that regulation and the effects thereof will be one of the major areas of research in fisheries economics over the next several years.[50]

8. Major policy issues II: The law of the sea and extended fisheries jurisdiction

The management of world fisheries is in the process of undergoing near revolutionary change as a consequence of the nine year Third United Nations Conference on the Law of the Sea which ended in December 1982. Prior to the

---

[50] Processing sectors of North American fishing industries are, more often than not, oligopsonistic in character. This raises the question of how the fishery would fare if the processing sector relevant to a given fishery were monopsonistic, rather than oligopsonistic. Would the fishery be managed as if it were controlled by a sole owner, i.e. would the common property aspects of the fishery vanish? What difference would it make if the harvesting sector responded to the monopsony by forming a monopoly? These questions were first raised by Crutchfield and Pontecorvo (1969) and were then examined in greater detail by Clark and Munro (1980), Munro (1982c) and Schworm (1983). The general conclusion is that a monopsonistic processing sector would not assure optimal management of the fishery. It does, however, hold forth the promise of better management than one currently enjoys with ineffective limited entry programs. Thus, one policy option would be to view the processing sector as a "natural monopsony" and act accordingly.

When first suggested, this policy proposal did seem rather far fetched, if not outlandish. Yet at the time of writing, there is a government sponsored move in Canada to turn the processing sectors of the fishing provinces of Newfoundland and Nova Scotia into virtual monopsonies [Canada, Department of Fisheries and Oceans (1983)].

Conference, coastal states,[51] such as the United States and Canada, had legal jurisdiction over fishery resources out to no more than 12 miles from shore. Fishery resources beyond 12 miles constituted international common property and were subject to management by international bodies[52] or were subject to no management whatsoever. As a consequence of the Conference, coastal states now have the right to establish so called Exclusive Economic Zones (EEZ) out to 200 miles from their shores. Within the EEZs coastal states have: "sovereign rights for the purpose of exploiting conserving and managing the natural resources, whether living or non-living...".[53] Thus, coastal states have virtual property rights to the fishery resources within their respective EEZ.

The significance of this change is that large, hitherto international, fishery resources have now come under coastal state control. Less than 10 percent of the world's fish harvests are taken in the remaining international waters [Eckert (1979, p. 116)].

A central argument on behalf of coastal state Extended Fisheries Jurisdiction (EFJ) was and is that the common problem cannot be addressed effectively when fisheries are international [Eckert (1979, ch. 5)]. Thus, for example, both the governments of the Untied States and Canada were subject to strong pressure to implement EFJ by their respective Atlantic coast fishermen who claimed that important fishery resources off their coasts were subject to heavy depletion by foreign, or so called distant water, fleets. There was indeed evidence that the fisheries were experiencing the classic, or Class I, common property problem. Fleets from over 15 countries competed for the resources [Munro and Pontecorvo (forthcoming)]. The fisheries were subject to regulation by an international body, which had as members the relevant coastal states and distant water nations, but it was argued that satisfactory management was impossible when the interests of so many diverse nations had to be reconciled. Control of individual coastal states brought with it the promise of superior management [Eckert (1979)].

EFJ was implemented on the promise that the hitherto international fisheries would be subject to superior management. If the full promise of EFJ is to be realized, however, coastal states will have to confront effectively three management issues raised by EFJ. The issues are:

(1) the intra-coastal state common property problem:
(2) the management of transboundary or shared stocks; and
(3) the establishment of co-operative arrangements with distant water nations.

The first issue arises from the fact that when a coastal state acquires a fishery resource through EFJ, the resource continues to be common property with respect

---

[51] States with ocean coastlines, as opposed to landlocked states.

[52] One example was the International Commission for the Northwest Atlantic Fisheries (ICNAF) which was responsible for the management of international fisheries off Canada and the United States from Greenland to the Carolinas.

[53] United Nations, Convention on the Law of the Sea, Article 56.

to coastal state fishermen. EFJ mitigates the common property problem; it does not eliminate the problem. The Class I or Class II common property problem can easily emerge with the EEZ.

An example of this danger is provided by Canada which gained control over heavily depleted fishery resources in the Northwest Atlantic. Canada succeeded in rebuilding the resources, but it also succeeded in transforming a Class I common property problem into a Class II problem. The very prospect of rebuilding fishery resources spurred an overexpansion in the coastal state harvesting *and* processing sectors [Munro and Pontecorvo (forthcoming)].

No new analysis, over and above what we have discussed in earlier sections, is required for the first issue. There is, however, one additional comment that should be made. If developed countries such as Canada, which enjoy substantial management capabilities, experience difficulties in effectively managing their newly acquired resources, then we must ask what the prospects are for the many developing coastal states?

The next issue, that of transboundary or shared stocks, is not a new one, but it has been greatly magnified by EFJ [Gulland (1980)]. Since fish are mobile many coastal states implementing EFJ find that some of the fishery resources they have acquired are, in fact, jointly owned by themselves and a neighboring coastal state or states. Fishery resources in the Gulf of Maine, formerly international, but now jointly owned by the United States and Canada, stand as an example.

Some theoretical work has now been done on this question in the form of extensions to the basic deterministic dynamic model, the discussion of which we have deliberately postponed to this time. We consider two examples. One is an article by Levhari and Mirman (1980) which examines the consequences of non-co-operation between joint owners of a transboundary resource. The second is an article by Munro (1979) which analyzes optimal management policies when the joint owners agree to co-operate, but when the joint owners' views as to the appropriate management strategy differ. Both articles confine themselves to the relatively simple case of two joint owners of a single fishery resource. Both employ dynamic game theory.

The Levhari–Mirman article has as its framework a Cournot–Nash model. Each joint owner takes the policy of the other owner as given and proceeds to attempt to maximize its net economic benefits, expressed as utilities, from the fishery. Equilibrium is achieved when each "player" believes that it cannot improve upon its situation, given the policies of the other. It is assumed, of course, that each joint owner exercises effective control over its respective fleets.

The Levhari–Mirman model is in discrete time and we have the following uncomplicated growth function:

$$x_{t+1} = x_t^{\alpha}, \tag{8.1}$$

where α is a constant, $0 < \alpha < 1$. The utility function with respect to current consumption of fish for each of the two joint resource owners, country 1 and country 2, is assumed to be given by

$$u_i(c_i) = \log c_i, \tag{8.2}$$

where c_i denotes current consumption of fish in country i. Finally, let β denote the common discount factor.[54]

Levhari and Mirman demonstrate that over an infinite time horizon, the optimal period by period consumption level in country 1 will approach $\bar{c}_1$ and in country 2 will approach $\bar{c}_2$ as $t \rightarrow \infty$, where

$$\bar{c}_1 = \bar{c}_2 = \left(\frac{1 - \alpha\beta}{2 - \alpha\beta} \right) x, \tag{8.3}$$

and that the steady state level of x, $x = \bar{x}$, will be

$$\bar{x} = \left(\frac{\alpha\beta}{2 - \alpha\beta} \right)^{\alpha/(1-\alpha)}. \tag{8.4}$$

They demonstrate as well that, if the two countries co-operate so as to maximize the discounted sum of the two countries' utilities, the steady state, or optimal, biomass will be given by

$$\hat{x} = (\alpha\beta)^{\alpha/(1-\alpha)}. \tag{8.5}$$

It is obvious that $\bar{x} < \hat{x}$.

We thus reach the following conclusions. Even if both countries solve the intra-coastal state common property problem and even if both have basically the same management goals, non-co-operation in the management of a transboundary resource will lead to overexploitation of the resource. Levhari and Mirman go on to show that in certain circumstances, non-co-operation will lead, not merely to overexploitation of the resource, but to outright extinction.

Generally speaking we can anticipate that co-operation will produce superior results to non-co-operation. If the joint owners are in agreement on their perception of optimal management policy – as is implicitly the case in the Levhari–Mirman example of co-operative management – the problem is relatively simple. The joint owners will have to bargain over the division of the benefits from the fishery, but the optimal management program will be the same as if there were a single owner.

Suppose, however, that the joint owners are not in agreement on their perception of the optimal management program. Then some form of compromise resource management program will have to be developed. This leads us to Munro's (1979) article on co-operative management of transboundary resources.

Munro considers three of the several reasons why the joint owners' perception of optimal management strategies might differ, these being different social rates of

---

[54] Levhari and Mirman do consider instances where the two countries have different discount factors. However, we are interested in the comparison between non-co-operative and co-operative steady state biomass levels. Here the authors confine themselves to the case of a common discount factor.

discount, different fishing effort costs and different consumer preferences. We shall confine ourselves to the first of the three.

Munro's model is a continuous time model and has as its foundation the Clark–Munro linear model. The accompanying game theoretic framework rests upon Nash's model of a two-person co-operative game [Nash (1953)].

Let δ_1 and δ_2 denote the social rates of discount of the joint owners of the resource, country 1 and country 2. By assumption $0 < \delta_1 < \delta_2 < \infty$. Fishing effort costs are assumed to be the same in both countries and both countries are assumed to sell into a world market where they face a perfectly elastic demand. Both countries have solved their intra-coastal state common property problem; each country seeks to maximize its own benefits from the fishery.

The difference in perceived optimal management policy is easy to describe. If we denote the optimal biomass levels as perceived by country 1 and country 2 as $x^*_{\delta_1}$ and $x^*_{\delta_2}$, respectively, we shall have $x^*_{\delta_1} > x^*_{\delta_2}$. Country 1, having a lower discount rate than country 2, will have a greater incentive to invest in the resource than its co-resource owner.

The procedure followed in determining an optimal compromise resource management policy is first to determine which of the feasible resource management policies are Pareto optimal. Of these Pareto optimal policies, one is then chosen through bargaining. (The theory of co-operative games is, of course, essentially a bargaining theory.) Bargaining determines how much weight is to be given to each joint owner's management preferences.

Prior to carrying out this two-stage procedure, it is necessary first to specify whether the co-operative games does or does not, allow side payments, i.e. transfers of wealth between the two "players", and whether the harvest shares are time variant. It is possible that political considerations will demand that the harvest shares be time invariant and be determined by prior negotiation.[55]

Suppose first that harvest shares are time invariant and that the co-operative game is without side payments. For there to be any compromise policy there must be at least one possible outcome that will cause both "players" to be at least no worse off than they would be without co-operation. Given that this condition can be met, Munro demonstrates that the compromise optimal biomass time path will be given by the following Golden Rule equation:

$$F'(x^*) - \frac{c'(x^*)F(x^*)}{p - c(x^*)} = \delta_3(t), \tag{8.6}$$

where $\lim_{t\to\infty}\delta_3(t) = \delta_1$. The discount rate $\delta_3(t)$ is a complex weighted average of δ_1 and δ_2, with the weights depending upon the relative bargaining strength of the two players or joint owners. Note that $\delta_3(t)$ is a function of time. Hence, $x^*_{\delta_3}(t)$ is also a function of time.

---

[55] In practice, it is common to determine harvest shares on the basis of historical catch records of the two countries.

The initial compromise optimal biomass, $x^*_{\delta_3}(0)$, will lie between $x^*_{\delta_1}$ and $x^*_{\delta_2}$. How close $x^*_{\delta_3}(0)$ is to $x^*_{\delta_2}$, as opposed to $x^*_{\delta_1}$, will depend, obviously, upon relative bargaining strengths. In any event, over time $x^*_{\delta_3}(t)$ will approach $x^*_{\delta_1}$ asymptotically. Thus the nature of the compromise is to give the management preferences of the high discount rate country relatively high weight in the present and near future and relatively low weight in the more distant future. The reason is quite simply that country 1 gives greater weight to the future than does its high discount rate co-resource owner.

Let us now jump to the case in which both the harvest shares are allowed to vary over time and in which side payments are assumed to be possible. The optimal compromise management policy is then straightforward. It becomes identical to that which would prevail if country 1 were the sole owner of the resource. Indeed, the optimal compromise calls for country 2, which places a lower value on the resource than its co-owner, to sell out entirely its claim to exploit the resource to country 1 at the commencement of the management program. The magnitude of country 1's payment to country 2 will once again depend upon relative bargaining strengths.

The outcome sounds extreme, yet it is interesting to note that at a meeting of Southeast Asian practitioners on problems created by EFJ, the possibility of a joint owner or owners of a transboundary resource leasing or selling out its (their) harvest rights to fellow joint owners was seriously considered [Christy (1978)]. Moreover, there exists a famous example in the Northeast Pacific in which joint owners did actually sell off their harvesting rights entirely to their co-owners.

The resource in question consisted of North Pacific fur seals. In the late nineteenth and early twentieth centuries, the resource was harvested by Russia and the United States on land (the Pribilof Islands) and at sea by Canada and Japan. The Russian and American harvesting operations were low cost; those of Canada and Japan were high cost. The theory suggests, not surprisingly, that an optimal co-operative management program with side payments would call for the high cost harvesters to sell out their harvesting rights to their low cost partners at the commencement of the program [Munro (1979)].

In 1911, in response to fears of overexploitation of the resource, an agreement, the North Pacific Fur Seal Agreement, was signed among the four harvesting countries. Under the agreement the high cost harvesters, Canada and Japan, were to reduce their harvesting activities to zero. The side payment was to take the form of Canada and Japan each being given an agreed upon share of the annual harvest, year after year, so long as the agreement remained in force. The agreement has been renewed many times and is in force in the present day. Currently, Canada and Japan each receive 15 percent of the annual harvest [Eckert (1979, p. 141)].

The third issue, the establishment of co-operative arrangements between coastal states and distant water nations, arises in the following manner. The "rules of the game" as they pertain to the fisheries aspects of EEZ management are laid down in the United Nations Law of the Sea Convention. Since the fisheries aspects of the Law of the Sea Conference were uncontroversial, coastal states appear willing to abide by these rules regardless of whether or not they have actually signed the Convention.

One aspect of the rules pertains to distant water nations that were exploiting many of the fishery resources now encompassed by EEZ.[56] Under the Convention, the coastal state is to establish TACs for each fishery resource in its EEZ. Any parts of the TACs that are surplus to coastal states harvesting capacity are to be made available to distant water nations. The coastal state is not, however, required to grant free access to these "surpluses". The coastal state is given the right to impose a wide range of terms and conditions, including fees, upon distant water fleet owners. Moreover, there is nothing to prevent the coastal state from working toward surpluses of zero over time.

A typical coastal state could not expect to eliminate all distant water fleets from its zone immediately upon the implementation of EFJ, unless it was prepared to incur exceedingly heavy surveillance and enforcement costs. It would, however, certainly be possible for the coastal state to phase out, through time, distant water fleet activity in its zone by imposing appropriate terms and conditions.

Several questions thus arise. First, is it in fact in the best interests of the coastal state to eliminate all distant water fleet activity in its zone? Or, is it, on the contrary, in the coastal states best interest to establish *co-operative fisheries arrangements* with distant water fleet owners so as to maintain their presence? A co-operative fisheries arrangement may be defined as any arrangement whereby a distant water fleet is allowed to participate in some or all aspects of the harvesting of a fishery resource within the coastal state EEZ or in the processing of the catch.[57]

Co-operative fisheries arrangements can usefully be divided into two broad categories: (a) joint venture type arrangements and (b) "fee fishing" type arrangements. In joint venture type arrangements, both distant water and coastal state entities engage in the harvesting/processing of the resources. Thus, for example, distant water vessels might harvest the resource and deliver the catch to onshore processing plants or coastal state vessels might harvest the resource and deliver the catch to foreign vessels with processing capacity. In "fee fishing" type arrangements, the distant water fleet both harvests the resource and processes the catch. The distant water fleet owner then pays the coastal state a fee or the equivalent thereof.

Secondly, if it is in the best interests of the coastal state to maintain a distant water presence in its zone, should the presence be short or long term. What form or forms of arrangement(s) are most likely to maximize the coastal state's benefits? If it is desirable to maintain a long-term distant water nation presence,

---

[56] United Nations, Convention on the Law of the Sea, Articles 61 and 62.

[57] There are deep sea fishing vessels that are equipped with processing capacity.

what conditions are necessary to ensure the willingness of distant water nations to participate in the coastal state fisheries on a long term basis.

There is much work yet to be done on this issue. A useful way to begin, however, is to cast the issue within the framework of international trade [Munro (1982b)]. Coastal states that enter into co-operative arrangements with distant water nations can be viewed as importing harvesting/processing services from these nations. Arguments from within the coastal state for considering distant water nation participation in coastal state fisheries are then to be seen as essentially free trade arguments. Conversely, arguments from within the coastal state for rejecting such participation are then to be seen as essentially arguments for protection, with the objects of protection being coastal state harvesting and/or processing interests.

The basic argument for considering distant water nation participation is simply that in certain coastal state fisheries distant water harvesters/processors may prove to have a comparative advantage *vis-à-vis* their coastal state counterparts. Therefore, the coastal state's net economic benefits from the fisheries will be maximized if it uses foreign, as opposed to domestic, harvesting/processing services [see Chan (1978)].

With respect to arguments for protection, one is concerned much less with protecting existing coastal state harvesting and processing activities than with giving these interests protection that will allow them to expand and replace distant water fleets. Hence, the most respectable of the protectionist arguments is a fisheries version of the infant industry argument [Munro (1982b)]. The infant industry argument is, of course, acknowledged by economists as a legitimate argument for temporary protection. Economists are, however, quick to add the warning that it is very difficult to determine *a priori* which infants do in fact have a reasonable opportunity of reaching maturity. If protection is given to an infant whose infancy proves to be permanent, then a burden may be imposed upon the economy that will also prove to be permanent.

If the coastal state should adopt protectionist policies having no economic justification, one consequence may be that some fishery resources will come to be "underexploited" or "underutilized". Consider the following example. Let it be supposed that the basic Clark–Munro linear model is applicable to a particular coastal state fishery. Let it be further supposed that a major issue in managing the fishery is whether the resource should be harvested by foreign or coastal state fleets. The cost of using foreign fishing effort is significantly less than the cost of using domestic fishing effort. If we denote the resource rent at any moment in time to be enjoyed by the coastal state when using domestic fishing effort as Π_D and the rent to be enjoyed when using foreign effort as Π_F, then it can be easily shown for any given biomass level and harvest rate that:

$$\begin{aligned} &\partial\Pi_D/\partial x \geq \partial\Pi_F/\partial x, \\ &\partial\Pi_D/\partial h < \partial\Pi_F/\partial x, \end{aligned} \tag{8.7}$$

where $\partial \Pi_D/\partial x = \partial \Pi_F/\partial x$ if and only if $\partial \Pi_D/\partial x = 0 = \partial \Pi_F/\partial x$. Thus, if harvesting costs are sensitive to biomass density, we shall have for any given biomass level:

$$\left\{ \frac{\partial \Pi_D/\partial x}{\partial \Pi_D/\partial h} > \frac{\partial \Pi_F/\partial x}{\partial \Pi_F/\partial h} \right\}. \tag{8.8}$$

The perceived marginal stock effect will be higher when domestic, as opposed to foreign, fishing effort is used.

From this it follows that, if x^* is the perceived optimal biomass level when coastal state effort is used and x^{**} the perceived optimal biomass level when distant water nation effort is used, then

$$x^* > x^{**}.$$

Hence, if the coastal state adopts protectionist policies and bars the use of foreign fishing effort, it may tend to "underutilize" the resource by virtue of having wilfully imposed upon itself higher than necessary exploitation costs.

Finally, with respect to the very much unexplored question of the co-operative arrangement terms and conditions necessary to ensure ongoing distant water nation willingness to participate in coastal state fisheries, we make one comment. The immediate consequence of EFJ for distant water nations was that it left them with large fleets in place, for which there were few alternative uses. Collectively they faced a large non-malleability of capital problem. Given this fact, they had an incentive to engage in co-operative fisheries arrangements with coastal states, even though the arrangements might be very short term and uncertain and even though the returns they earned might have done little more than cover their operating costs.

The situation is, however, obviously transitory. In time, the existing fleets will reach the end of their economic lives. Whether distant water fleet owners, be they private or state, will be prepared to reinvest, even partially, in fleet capacity remains an open question [Munro (1983)]. If the returns on such reinvestment (political as well as economic) appear to be low and/or uncertain, then the reinvestments will be unlikely to occur and distant water participation in coastal state EEZ could contract severely, if not disappear. Unequivocal underutilization of the major fishery resources could then be the consequence.

9. Conclusions

The importance of, and interest in, the economics of fisheries management has grown substantially over the past fifteen years as a consequence of the United Nations Third Conference on the Law of the Sea. The Conference has brought about a near revolution in the management of world fisheries.

Nonetheless, the central economic problem in the management of fisheries remains the same as it was at the dawn of modern fisheries economics thirty years ago. Fishery resources constitute common property resources with the consequence that fisheries are subject to market failure. Unless a regime of judicious and effective government regulations is implemented, the common property aspect of the fisheries will result, either in excessive depletion of the resource, or in the dissipation of potential economic benefits from the fishery through the emergence of redundant labor and capital.

The major development in fisheries economics that we have chosen to stress in this chapter is the shift away from static to dynamic or capital-theoretic analysis. Dynamic considerations in turn lead naturally to a consideration of problems arising from uncertainty.

The shift to capital-theoretic analysis arises from the recognition that fishery resources, like all other natural resources, constitute capital assets in that they are capable of yielding a stream of benefits to society through time. The recognition is not recent. Indeed, it has been there since the origin of modern fisheries economics, as the work of Gordon (1956) and of Scott (1955) make evident. The late arrival of capital-theoretic analysis in fisheries economics is to be explained by the fact that economists had to await the development of adequate mathematical techniques.

Four areas in fisheries economics are bound to be given emphasis in future years. The first three arise from the problems of implementation of EFJ, the regulation of fisheries and uncertainty. The final area consists of the empirical estimation and testing of fisheries models, an area in which development to date has been limited[58] and in which the scope for future research is great.

References

Andersen, Peder and Jon G. Sutinen (forthcoming) "A Survey of Stochastic Bioeconomics", *Marine Resource Economics*.

Anderson, Lee G. (1977) *The Economics of Fisheries Management* (Johns Hopkins, Baltimore).

Baker, D. (1980) *The Capital Development Fund: A Capital Assistance Plan for Fishermen*, Report prepared for the Government of Canada, Department of Fisheries and Oceans, Halifax.

Bishop, Richard C. and Karl C. Samples (1980) "Sport and Commercial Fishing Conflicts: A Theoretical Analysis", *Journal of Environment Economics and Management* 7, 220–233.

Brown, Gardner J. (1974) "An Optimal Program for Managing Common Property Resources with Congestion Externalities", *Journal of Political Economy* 82, 163–174.

Canada, Commision on Pacific Fisheries Policy (1982) *Final Report*.

---

[58] For a discussion of the problems involved in estimating such models, see Uhler (1980). Having said this, however, it is important to acknowledge that some empirical work has been carried out successfully. As an example, one can point to a recently completed massive econometric study of the Newfoundland groundfishery, which encompasses, not only the harvesting sector, but the processing and demand sectors as well. See Schrank et al. (1984).

Canada, Department of the Environment, Fisheries and Marine Service (1976) *Policy for Canada's Commercial Fisheries* (Ottawa).

Canada, Department of Fisheries and Oceans (1983) Remarks by the Honourable Pierre De Bané, Minister of Fisheries and Oceans in the House of Commons, 14 November 1983 (Ottawa).

Canada, Task Force on Atlantic Fisheries (1982) *Report.*

Chan, Kenneth Shun-Yuen (1978) "The Economic Consequences of the 200-Mile Seabed Zone", *Canadian Journal of Economics* 11, 314–318.

Charles, Anthony T. (1983) "Optimal Fisheries Investment Under Uncertainty", *Canadian Journal of Fisheries and Aquatic Science* 40, 2080–2091.

Christy, Francis T. Jr. (1973) *Fishermen's Quotas: A Tentative Suggestion for Domestic Management*, University of Rhode Island, Law of the Sea Institute, Occasional Paper No. 19.

Christy, Francis T. Jr. (Ed.) (1978) *Law of the Sea: Problems of Conflict and Management in Southeast Asia*, International Center for Living Aquatic Resources Management and Institute of Southeast Asian Studies, ICLARM Conference Proceedings No. 2.

Christy, Francis T. Jr. (1982) *Territorial Use Rights in Marine Fisheries: Definitions and Conditions*, Food and Agriculture Organization of the United Nations, Fisheries Technical Paper No. 227.

Clark, Colin W. (1976) *Mathematical Bioeconomics: The Optimal Management of Renewable Resources* (Wiley, New York).

Clark, Colin W. (1980) "Towards a Predictive Model for the Regulation of Fisheries", *Canadian Journal of Fisheries and Aquatic Science* 37, 1111–1129.

Clark, Colin W. (1981) "The Economics of Whaling: A Two-Species Model", *Simulation.*

Clark, Colin W. (1982) "Models of Fishery Regulation", in: L.J. Mirman and D.F. Spulber (eds.), *Essays in the Economics of Renewable Resources* (North-Holland, Amsterdam) pp. 273–287.

Clark, Colin W. and Geoffrey P. Kirkwood (1979) "Bioeconomic Model of the Gulf of Carpentaria Prawn Fishery", *Journal of the Fisheries Research Board of Canada* 36, 1304–1312.

Clark, Colin W. and Gordon R. Munro (1975) "The Economics of Fishing and Modern Capital Theory: A Simplified Approach", *Journal of Environmental Economics and Management* 2, 92–106.

Clark, Colin W. and Gordon R. Munro (1980) "Fisheries and the Processing Sector: Some Implications for Management Policy", *The Bell Journal of Economics* pp. 603–616.

Clark, Colin W., Frank H. Clarke and Gordon R. Munro (1979) "The Optimal Exploitation of Renewable Resource Stocks: Problems of Irreversible Investment", *Econometrica* 47, 25–47.

Clarke, Frank H. (forthcoming) "Applying Abstract Control Theory to Concrete Problems: A Casebook", in: R. McKelvey (ed.), *A Short Course on Environmental and Natural Resource Mathematics* (American Mathematical Society, Providence).

Copes, P. (1972) "Factor Rents, Sole Ownership and the Optimal Level of Fisheries Exploitation", *Manchester School of Economic and Social Studies* 40, 145–164.

Crutchfield, James (1956) "Common Property Resources and Factor Allocation", *Canadian Journal of Economics and Political Science* 22, 292–300.

Crutchfield, James (1982) *The Pacific Halibut Fishery* Economic Council of Canada, Technical Report No. 17, The Public Regulation of Commercial Fisheries in Canada.

Crutchfield, James and Giulio Pontecorvo (1969) *The Pacific Salmon Fisheries: A Study of Irrational Conservation* (Johns Hopkins, Baltimore).

Crutchfield, James and Arnold Zellner (1962) *Economic Aspects of the Pacific Halibut Fishery*, U.S. Department of the Interior, Fishery Industrial Research.

Dasgupta, Partha (1982) *The Control of Resources* (Harvard University Press, Cambridge).

Dasgupta, Partha and G.M. Heal (1979) *Economic Theory and Exhaustible Resources* (Cambridge University Press).

Doubleday, W.G. (1976) "Environmental Fluctuations and Fisheries Management", International Commission for the Northwest Atlantic Fisheries, *Selected Papers Number 1, 1976*, pp. 141–150.

Eaton, B. Curtis and Richard G. Lipsey (1981) "Capital Commitment and Entry Equilibrium", *The Bell Journal of Economics* 2, 596–604.

Eckert, Ross D. (1979) *The Enclosure of Ocean Resources: Economics and the Law of the Sea* (Hoover Institution Press, Stanford).

Fraser, G.A. (1979) "Limited Entry: Experience of the British Columbia Salmon Fishery", *Journal of the Fisheries Research Board of Canada* 36, 754–763.

Gause, G.F. (1935) *La Théorie Mathematique de la Lutte Pour La Vie* (Hermann, Paris).

Gordon, H. Scott (1954) "The Economic Theory of a Common Property Resource: The Fishery", *Journal of Political Economy* 62, 124–142.

Gordon, H. Scott (1956) "Obstacles to Agreement on Control in the Fishing Industry", in: Ralph Turvey and Jack Wiseman (eds.) *The Economics of Fisheries* (Food and Agriculture Organization of the United Nations, Rome) pp. 65–72.

Gulland, J.A. (1980) *Some Problems of the Management of Shared Stocks* Food and Agriculture Organization of the United Nations, Fisheries Technical Paper No. 206.

Hannesson, R. (1975) "Fishing Dynamics: A North Atlantic Cod Fishery", *Canadian Journal of Economics* 8, 151–173.

Huppert, D.D. (1979) "Implications of Multipurpose Fleets and Mixed Stocks for Control Purposes", *Journal of the Fisheries Research Board of Canada* 36, 845–854.

Klein, Benjamin and K.B. Leffler (1981) "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy* 89, 615–641.

Kurz, M. (1968) "Optimal Economic Growth and Wealth Effects", *International Economic Review* 9, 348–357.

Larkin, P.A. (1977) "An Epitaph for the Concept of Maximum Sustained Yield", *Transactions of the American Fisheries Society* 106, 1–11.

Levhari, D. and L.J. Mirman (1980), "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution", *The Bell Journal of Economics* 11, 322–334.

Levhari, D., R. Michiner and L.J. Mirman (1981) "Dynamic Programming Models of Fisheries", *American Economic Review* 71, 649–661.

Lewis, Tracy R. (1981) "Exploitation of a Renewable Resource Under Uncertainty", *Canadian Journal of Economics* 14, 422–439.

Lewis, Tracy R. and R. Schmalensee (1979) "Non-Convexity and Optimal Harvesting Strategies for Renewable Resources", *Canadian Journal of Economics*, 12, 677–691.

Long, N.V. (1977) "Optimal Exploitation and Replenishment of a Natural Resource", in: John Pitchford and Steven Turnovsky (eds.) *Application of Control Theory to Economic Analysis* (North-Holland, Amsterdam).

McKelvey, Robert (1983) "The Dynamics of Open-Access Exploitation of a Renewable Resource: The Case of Irreversible Investment", Department of Mathematics, University of Montana, Interdisciplinary Series No. 21.

May, Robert M., John R. Beddington, Colin W. Clark, Sidney J. Holt and Richard M. Laws (1979) "Management of Multispecies Fisheries", *Science* 205, 267–277.

Mendelssohn, R. (1979) "Determining the Best Trade-Off Between Expected Economic Return and the Risk of Undesirable Events When Managing a Randomly Varying Population", *Journal of the Fisheries Research Board of Canada* 36, no. 8, 939–947.

Mitchell, Carlyle L. (1978) "Stock Adjustment Models: Canada's East Coast Groundfish Fishery", Ph.D. thesis, Department of Economics, University of Ottawa.

Moloney, D.G. and P.H. Pearse (1979) "Quantitative Rights as an Instrument for Regulating Commercial Fisheries", *Journal of the Fisheries Research Board of Canada* 36, 859–866.

Munro, Gordon R. (1979) "The Optimal Management of Transboundary Renewable Resources", *Canadian Journal of Economics* 12, 355–376.

Munro, Gordon R. (1981) "The Economics of Fishing, an Introduction", in: John A. Butlin (ed.), *Economics and Resources Policy* (Longman, London) pp. 129–141.

Munro, Gordon R. (1982a) "Fisheries, Extended Jurisdiction and the Economics of Common Property Resources", *Canadian Journal of Economics* 15, 405–425.

Munro, Gordon R. (1982b) "Co-operative Fisheries Arrangements between Pacific Coastal States and Distant-Water Nations", in: H.E. English and A.D. Scott (eds.) *Renewable Resources in the Pacific*, Proceedings of the 12th Pacific Trade and Development Conference (International Development Research Centre, Ottawa) pp. 247–254.

Munro, Gordon R. (1982c) "Bilateral Monopoly in Fisheries and Optimal Management Policy", in: L.J. Mirman and D.F. Spulber (eds.) *Essays in the Economics of Renewable Resources* (North-Holland, Amsterdam) pp. 187–202.

Munro, Gordon R. (1983) "Foreign Access to EEZs and the Derivation of Coastal State Benefits: Methods and Techniques", in: Food and Agriculture Organization of the United Nations *Report of the Expert Consultation on the Conditions of Access to the Fish Resources of the Exclusive Economic*

Zones, FAO Fisheries Report No. 293 (Rome) pp. 143–151.

Munro, Gordon R. and Giulio Pontecorvo (forthcoming) "The Taking of Living Resources: Expectations and Reality", *Proceedings of the Seventeenth Annual Conference of the Law of the Sea Institute* Oslo, 1983.

Nash, J.F. (1953) "Two Person Co-operative Games", *Econometrica* 21, 128–140.

Neher, P.A. (1974) "Notes on the Volterra Quadratic Fishery", *Journal of Economic Theory* 6, 39–49.

Paterson, Frederick M. and Anthony C. Fisher (1977) "The Exploitation of Extractive Resources: A Survey", *The Economic Journal* 87, 681–721.

Plourde, G.C. (1970) "A Simple Model of Replenishible Resource Exploitation", *American Economic Review* 60, 518–522.

Plourde, G.C. (1971) "Exploitation of Common-Property Replenishible Resources", *Western Economic Journal* 9, 256–266.

Quirk, J.P. and V.L. Smith (1970) "Dynamic Economic Models of Fishing", in: A.D. Scott (ed.), *Economics of Fisheries Management: A Symposium* (University of British Columbia, Institute of Animal Resource Ecology, Vancouver) pp. 3–32.

Reed, William J. (1979) "Optimal Escapement Levels in Stochastic and Deterministic Harvesting Models", *Journal of Environmental Economics and Management* 6, 350–353.

Schaefer, M.B. (1957) "Some Considerations of Population Dynamics and Economics in Relation to the Management of Marine Fisheries", *Journal of the Fisheries Research Board of Canada* 14, 669–681.

Schaefer, M.B. and R.J.H. Beverton (1963) "Fishery Dynamics – Their Analysis and Interpretation", in: M.N. Hill (ed.) *The Sea*, vol. 2, (Interscience, New York), 464–483.

Schrank, William E., Eugene Tsoa and Noel Roy (1984), *An Econometric Study of the Newfoundland Groundfishery: Estimation and Simulation*. Study prepared for Fisheries and Oceans Canada (Ottawa).

Schworm, William E. (1983) "Monopsonistic Control of a Common Property Renewable Resource", *Canadian Journal of Economics* 16, 275–287.

Scott, Anthony D. (1955) "The Fishery: The Objectives of Sole-Ownership", *Journal of Political Economy* 63, 116–124.

Scott, Anthony D. (forthcoming) "Catch Quotas and Shares in the Fishstock as Property Rights", in: E. Miles and R. Stokes (eds.), *Essays in Honor of James Crutchfield* (University of Washington Press, Seattle).

Scott, Anthony D. and Philip A. Neher (1981) *The Public Regulation of Commercial Fisheries in Canada* (Economic Council of Canada, Ottawa).

Smith, V.L. (1977) "Control Theory Applied to Natural and Environmental Resources: An Exposition", *Journal of Environmental Economics and Management* 4, 1–24.

Spulber, Daniel F. (forthcoming) "Fisheries and Uncertainty", in: A.D. Scott (ed.) *Progress in Natural Resource Economics* (Oxford University Press, Oxford).

Strand, Ivar E. Jr. and Virgil J. Norton (1980) "Some Advantages of Landings Taxes in Fishery Management", in: John H. Grover (ed.) *Allocation of Fishery Resources* (Food and Agriculture Organization of the United Nations, Rome) 411–417.

Uhler, R.S. (1980) "Least Squares Regression Estimates of the Schaefer Production Model: Some Monte Carlo Results", *Candian Journal of Fisheries and Aquatic Science* 37, 1284–1294.

United Nations (1982) "Convention on the Law of the Sea", Third Conference on the Law of the Sea.

United States Congress, "Fishery Conservation and Management Act of 1976", PL 94-265.

Walters, C.J. and R. Hilborn (1978) "Ecological Optimization and Adaptive Management", *Annual Review of Ecological Systems* 9, 157–188.

Williamson, Oliver E. (1983) "Credible Commitments: Using Hostages to Support Exchange", *American Economic Review* 83, 519–540.

Chapter 15

THE ECONOMICS OF OUTDOOR RECREATION

KENNETH E. McCONNELL

University of Maryland

1. Introduction

The economics of outdoor recreation deals with the supply of and demand for natural resources for recreational purposes. This chapter will survey conceptual and empirical approaches, problems, and solutions encountered in applying economics to the provision of such natural resources. It will show how the evolution of the economics of outdoor recreation was influenced by the distinctive nature of markets for outdoor recreation.

The economics of outdoor recreation is a more narrow subject than might be conjured up by the phrase "outdoor recreation". This branch of the economics of natural resources focuses on the use of substantial resources such as forests, rivers, canyons, or bays for recreational purposes. Thus, while activities such as tennis, golf, and polo fall within the rubric of outdoor recreation as popularly conceived, they are of only passing interest for this paper because they do not have intensive requirements for natural resources. The notion of the intensiveness of resource use as a criterion for considering activities to study is not highly selective, and should be interpreted loosely.

The economics of outdoor recreation has developed from two distinct stimuli. The first stimulus was the development of substantial government holdings of land and the heavy subsidization of agricultural water projects, both in the western United States. Substantial government involvement in the exploitation of natural resources, especially in agriculture, required some accountability. Thus it was natural that intellectual resources, especially from land grant institutions, would examine the ramifications of federal land and water resource policies, particularly in the West. These policies included the formation of national parks and forests, range management, dam construction, and irrigation subsidies. This development makes the economics of outdoor recreation a progeny of land resources economics, and in its role here it retains a positive nature, with more

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

concern on the sources and uses of data and less concern with the structure of models.

The second stimulus is the more recent emergence of the environmental movement. This movement developed from a consciousness about the value of the nation's environmental assets. The commitment to maintain and improve the quality of these assets is, in part, motivated by the belief that better recreational opportunities will be created. For example, one of the goals of the program to curtail water pollution is to develop fishable, swimmable waters. As a consequence, a significant portion of outdoor recreation economics has grown as a tool to evaluate the benefit of environmental improvements. In this role, the economics of outdoor recreation has developed as an offspring of applied welfare economics, with emphasis on the structure of choice from which decision models are derived.

Outdoor recreation economics is an imperfect blend of responses to both stimuli. Attempts to use sophisticated choice models tend to conflict with the limitations of survey data. But the resolution of these conflicts often results in the improvement of survey design for future use. Empirical results, consistently at odds with theoretical models, lead to the revision of models, and a better understanding of individual behavior. While this process leads to frequent revisions of and challenges to accepted methods, it has produced some results which are useful in the valuation of the recreational use of natural resources. This process and the results will be the topic of this survey.

The survey will be divided into three parts. In the first part (Section 2), the extent of and rationale for government involvement in the provision of outdoor recreation services will be explored. The second part (Section 3) will investigate conceptual and empirical issues in estimating the demand or benefits of outdoor recreation. In the third part (Section 4), some issues relating to the quality of outdoor recreation will be explored. Reflecting the allocation of current research effort, the greatest proportion of the chapter (Section 3) will be devoted to empirical and conceptual demand issues.

A bird's eye view of the area will help orient the reader. Outdoor recreation is a service produced and consumed by an individual in conjunction with a natural resource. The chief characteristic of the production of the service, from the individual's perspective, is that he must transport himself to the site to consume the service, rather than have the commodity itself transported. For the individual, the scarce resources such as time, transportation services, etc. are central to the decision process of whether and how much to use a particular resource. Thus, the formation of the demand for outdoor recreation services requires focusing on the allocation decisions for household resources. For society, the chief characteristic of outdoor recreation is the natural resource requirements. The economics of outdoor recreation, when dealing with the supply side, focuses on the opportunity costs of natural resource facilities when used to produce recreational services.

2. Government involvement in the provision of outdoor recreation services

There is a large body of literature in the economics of natural resources dealing with the efficiency of government intervention in natural resource markets when markets fail to provide information on uncertainty, irreversibility, or externalities. Of this considerable literature, very little asks why, in principle, the public sector should provide natural resources on which recreation takes place. Perhaps the relative scarcity of research on the welfare aspects of government provision of resources for recreation activities simply recognizes a fact: governments do provide a great many natural resources for the primary purpose of outdoor recreation. Hence, no rationale is needed. Nevertheless, it is instructive to ask why the government is providing these natural resources, and perhaps to gain some understanding of the circumstances in which it might be in society's best interest to have these services provided by the private sector.

The welfare aspects of providing facilities for outdoor recreation are explicitly considered in Robinson (1967), Clawson and Knetsch (1966), and Davidson, Adams and Seneca (1966). There is no clear consensus among these works about the rationale for government intervention. Robinson suggests that recreation resources are provided primarily as merit wants. Clawson and Knetsch argue that a combination of economics of scale and externalities provide the rationale for government intervention, while Davidson, Adams and Seneca reason that a mix of public goods and externalities prevents the market from being efficient.

While the economics of natural resources is replete with debate over government intervention and market failure, we are interested only in the following narrow question: What are the circumstances which promote inefficiency in the market's utilization of natural resources producing recreational services?

The answer relies on the nature of outdoor recreation. Imagine a relatively large natural resource, say a national forest. Recreation services are produced on this resource when individuals utilize the resource for their own pleasure. The distinctive feature of outdoor recreation services is that services are produced by the individuals who consume them. For example, in camping, an individual takes his own time, tent, sleeping bag, and supplies, uses only the natural resources, and produces camping services which he also consumes. The supply decisions exogenous to the individual involve only having the natural resource available in a relatively natural state. Barring congestion and environmental degradation, there are no variable costs of production. Hence, for the resource owner, the marginal cost of providing the recreational services is zero, on the assumption that the individual produces no externalities.

Now, suppose that this resource is to be privately owned and used to produce recreation services. It seems clear that a monopolistic market structure will develop. For the monopolistic market structure, the case for inefficiency is well known here. Two possibilities exist. In Figure 15.1, *d* is the relevant aggregate

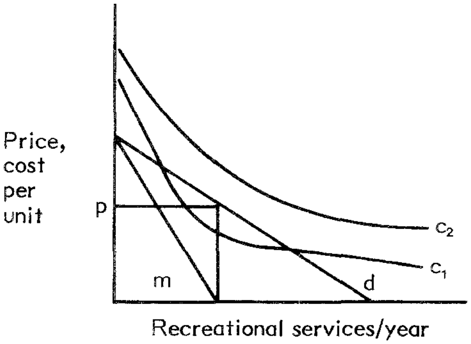

Figure 15.1. Cost and demand functions for recreational services.

demand schedule and m the marginal revenue schedule. The only costs are fixed costs, which are composed solely of opportunity costs. That is, the only fixed cost of maintaining a resource in its natural state is the return it would receive in its best alternative case. In the first case, where average fixed costs are c_1, price p will be set where marginal revenue is zero. This case is inefficient, because the price exceeds the social marginal cost, which is zero. However, some services will be provided because the price exceeds the average (fixed) costs. In the second case, where average fixed costs are c_2, price is less than the average costs, so that no recreational services will be produced. The resource will be used for other productive purposes. This outcome, too, will be inefficient if the area under the demand curve, a measure of the total worth of the resource in producing recreation services, exceeds the total cost of keeping the resource producing recreational services, its fixed costs.

The possibility of a competitive outcome deserves a word. With no capacity constraint and with zero marginal cost, the profit maximizing firm will increase output of service indefinitely. This is a logical impossibility. Some combination of capacity constraints, positive marginal cost, or declining marginal revenues must be induced to produce finite output. Since the scale of many natural resources is so large, it seems unlikely that in the relevant range average cost will rise. A long-run competitive solution, while logically possible, seems quite unlikely.

Thus, the primary rationale for government involvement in the provision of natural resources for outdoor recreation services is the economies of scale argument. Resources which are relinquished to the private sector will provide less than optimal levels of service, or will provide none at all. There are, of course, other logical planks in the intervention platform. For example, recreational services tend to be one of several types of services produced simultaneously by natural resources. The multiple use aspects of National Forests have received much attention. However, there appears to be no compelling reason why the multiple outputs argument necessarily leads to government intervention. As long as benefits can be appropriated and users made to pay for their satisfaction, private enterprise can be efficient in providing resources for multiple use.

Irrespective of the economic rationale for government provision of resources for outdoor recreation, some simple descriptive material makes clear the extent to which this provision takes place. Two sorts of measures are available. One type of measure gauges the quantity of services demanded. This measure is reported in units of numbers of participants and numbers of days or visits. For example, for hunting, 32 percent of hunting days in 1980 took place on publicly-owned land [U.S. Department of Interior (1982, p. 80, table 31)]. For western regions of the United States, this measure was much higher: 63 percent for the Rocky Mountain region and 54 percent for Pacific states. From the same study, we learn that 75.4 percent of participants in nonconsumptive wildlife recreation [bird watching, wildlife photography, etc. (p. 107, table 52)] visited a public area of one sort or another. Similar descriptive measures are available for activities such as camping and hiking. Other activities, such as all saltwater fishing and much freshwater fishing, by their nature, take place only on publicly-owned or controlled resources.

Another measure reflecting the public's provision of natural resources is the quantity of publicly-owned land. In total, about one-third of all the land area in the United States is federally owned. The vast majority of this land is in the West. In Alaska, 98.3 percent of all lands are federally owned. In Nevada, the federal government owns 86.1 percent of the land. Over 90 percent of the federal government's 775 million acres is used for forest and wildlife, grazing, and parks and historic sites, and is quite suitable for recreational use. Thus, we see that there are large areas of undeveloped natural resources which serve as the basic public input into the production of private outdoor recreation services.

Many individual facilities which provide access to large resources are also publicly owned. Though statistics are hard to come by, it is clear that federal, state, and local governments provide operating and capital costs for boat launching ramps, campgrounds, nature trails, picnic grounds, tennis courts, and many other combinations of resource and structure. Agencies such as the Coast Guard and the Forest Service, not initially created to serve recreational interests, now spend considerable effort to do so. Thus, while there is some private provision of recreational services, it is dwarfed in comparison with the public sector.

3. The demand for and value of outdoor recreation facilities

Governments, federal, state, and local, are committed to providing outdoor recreation sites and facilities. The provision of these facilities and resources requires government agencies to deal with management issues. These issues range from simple benefit–cost analysis to complex questions of optimal mix of recreation services with other kinds of output. The resolution of most management questions, in principle at least, requires knowing the demand for outdoor recreation services. Thus, it is not surprising that the greatest quantity of work in the economics of outdoor recreation involves demand issues. It is quite rare to find research on the economics of outdoor recreation which does not focus primarily on the demand site. This section of the chapter deals with the approaches and problems of estimating the demand for outdoor recreation.

There are basically three approaches to estimating the demand for and value of outdoor recreation. They are: (1) the travel cost approach; (2) contingent valuation approaches; and (3) variants and combinations of the household production function–hedonic price approach. In addition, there are reduced form models used for forecasting. The following section will attempt to show how these approaches are derived from theory and used in practice. Since they attempt to measure the same phenomena, they presumably have some common conceptual ground. Part of the explication of these approaches will be devoted to exploring conceptual and practical similarities and differences. (Additional discussion of these methods can be found in Chapter 6 of this Handbook.)

To begin the analysis, we express the benefits of a natural resource in recreational use as it relates to the demand for outdoor recreation. Then we show how various applied research problems arise in trying to estimate demand. Suppose the natural resource under evaluation is a recognizable resource, confined to a locality. The reasoning works better when the resource has only one point of entry, though this is not necessary. Let

$$x = f(p, w) \tag{1}$$

be the representative user's demand for trips or visits x of specified length to the site, where w is a vector of exogenous variables and p is the cost of entry. Suppose that the area around the site is divided into m different regions or zones. The regions, indexed by i, may differ because of distance from the site, population composition, or the opportunities facing recreation users. The per user demand in region i is

$$x_i = f(p_i, w_i), \tag{2}$$

and aggregate demand in region i is

$$n_i x_i = n_i f(p_i, w_i), \tag{3}$$

where n_i is the number of users in the ith region.

The total benefits to consumers of operating the site as a recreational resource can then be computed as the area under the demand curve for each user, aggregated across users for each region, and then aggregated across all zones.

Then aggregate benefits to recreational users are given by

$$B = \sum_{i=1}^{m} n_i \int_{p_i}^{p_i^*} f(p, w_i)\, \mathrm{d}p, \tag{4}$$

where p_i is the prevailing cost of attending the site from the ith zone and p_i^* is the reservation price $[f(p_i^*, w_i) = 0]$ for the ith distance zone.

In defining net benefits, we are forced to make explicit the implicit assumption in most of recreation economics: for purposes of welfare measurement, Marshallian demand functions may be substituted for Hicksian demand functions without causing grave errors. With this assumption, the benefit measures in (4) can be described as the amount of money users would pay for the right of access to the site. While Willig (1976) shows that the magnitude of error in this approximation is likely to be small, it is a troublesome approximation for several reasons [see Bockstael and McConnell (1979), Hanemann (1980), and Hausman (1982), and the discussion in Chapter 6 of this Handbook]. First, we are implicitly dealing with large price changes, large enough to price everyone out of the market. Second, there are many cases where we are dealing with fairly unique natural resources, so that a large welfare effect from eliminating the resource may be plausible. Thus, we should not forget the assumptions which allow us to use the Marshallian triangle as willingness to pay for the right to access.

Having all the relationships which go into computing benefits as in (4) would allow one to answer the practical questions which have been addressed to economists over the last several decades. Note that computation of total benefits requires the following kinds of information or predictions:

(1) the total number of users;
(2) trips per user;
(3) how trips per user respond to site price increases;
(4) how trips per user respond to exogenous variables;
(5) the reservation price; and
(6) the number of regions exercising effective demand.

The following sections deal with the various approaches to estimating these relationships and computing benefits as in (4). (Additional discussion of the travel cost method is found in Chapter 6 of this Handbook.)

3.1. The travel-cost method

The travel-cost method is an approach which uses variations in travel costs to a recreational site to estimate the demand for one site. The intellectual origin of the travel cost method is remarkable. It stems from a letter by Harold Hotelling written in response to a solicitation by the National Park Service. The solicitation was sent to ten economists, asking them to suggest methods for measuring, for national parks and other areas, "economic benefits derived from these areas in excess of the economic returns and benefits that would accrue if the areas were used for other purposes" (letters from A.E. Demaray in National Park Service, 1949, p. 2). Hotelling responded with a description of the essence of what has come to be known as the travel-cost method.

Because travel costs are an important determinant in visiting a site, variations in travel cost help in demand estimation. This phenomenon is put aptly by Burt and Brewer (1971):

> ...the consumer is transported to the commodity for consumption to take place, rather than vice versa. This attribute of outdoor recreation is advantageous for statistical estimation of demand equations because the costs that must be incurred to consume the recreational services provide surrogate prices with more variation in a sample than would usually be generated by market phenomena observed either over time or over space (p. 813).

This variation in the costs of reaching a site has been the basis for travel-cost models since the earliest applications. Some examples give insight into the development of the method.

One of the earliest studies to use the travel-cost method is Clawson's (1959) work. This study of Yosemite, Grand Canyon, Glacier, and Shenandoah National Parks does not estimate per capita demand functions, but it does plot visit rates measured as participants per 100 000 population against costs per visit. In addition, Clawson uses a free hand fit of the plot to predict the impact of increasing entrance fees at each of the parks. These points are plotted to give the aggregate demand curve, or what Clawson called "the true demand curve for the recreation opportunity itself" (p. 26). Though Clawson does not calculate consumers' surplus for any national park, he indicates clearly how the aggregate demand curves could be used for that purpose.

In a study of the Kerr Reservoir in North Carolina, Knetsch (1964) estimates the following demand curve, in natural logs,

$$\ln(v + 0.8) = 3.82 - 2.39 \ln c$$

for 12 distance zones, where "v is the visit rate per thousand population per zone of origin and c is the dollar cost of travel" (p. 1151). Knetsch simulates entrance price increases by raising c, computing the visit rate per thousand population for each zone, aggregating each zone by multiplying by the relative population figure in each zone, and adding up across zones. The plot of each simulated price and aggregate quantity combination is the aggregate demand curve. Then Knetsch computes the "recreation benefit estimates for Kerr Reservoir, based on the integral of the imputed demand curve... [to be] about $1.6 million" (p. 1154), a clearly derived estimate of the willingness to pay for access to a recreational site.

Brown, Singh and Castle (1964) explore the travel-cost approach and develop a working model of benefit estimation for fishing in Oregon for salmon-steelhead. The state of Oregon is divided into 35 zones. Using a sample frame of licensed anglers, Brown et al. mailed questionnaires to determine the number of days fishing in each of the 35 zones. With the information on the total number of days per zone, as well as reported information on expenditures, family income, and distance from the site, they estimate the following relationship:

$$v_i/P_i = h(c_i, d_i, y_i), \tag{5}$$

where

v_i = days of fishing from ith zone,
P_i = population of the ith zone,
c_i = expenditures per day in the ith zone,
d_i = distance from the site to the ith zone, and
y_i = family income of the ith zone.

The following equation is reported by Brown et al.:

$$\ln v_i/P_i = 0.951 - \underset{(0.029)}{0.128}\, c_i - \underset{(0.002)}{0.002}\, d_i + \underset{(0.002)}{0.007}\, y_i, \tag{6}$$

where the standard errors are reported below the coefficients. Aggregate demand in zone i is

$$v_i = P_i h(c_i, d_i, y_i), \tag{7}$$

where $v_i = n_i x_i$.

Travel-cost models occasionally estimate per capita demand functions. Thus, per capita functions, such as $h(c, d, y)$ imply two types of decisions: the decision about how many days to enjoy and the decision about whether to participate. Consumers' surplus for the site can be calculated from eqs. (6) and (7) as

$$B = \sum_{i=1}^{m} P_i \int_0^{p_i^*} h(p + c_i, d_i, y_i)\, \mathrm{d}p, \tag{8}$$

where, in this case, $m = 35$, $p_i^* \to \infty$, and P_i, c_i, d_i, and y_i are gathered from the survey.

Several assumptions about entrance fees in the travel cost approach are revealed by examining (8). First, the fundamental assumption of this method is obvious: forming the argument $p + c_i$ assumes that individuals respond the same to dollars spent on reaching the site, regardless of how the dollars are spent. A dollar spent on a site entrance fee has the same impact on quantity demanded as a dollar spent on, say, gasoline, by assumption. Second, the formulation in (8) assumes that the prevailing entrance fee is zero, or if it is not zero, the entrance fee collected is not counted as part of the benefits from using the site.

From expression (8), we see that the value of a site to consumers from the travel-cost approach is the sum of the per capita value, aggregated across population within a zone, and then aggregated across zones. We can arrive at the same measure another way. Expression (7) gives the aggregate demand from zone i at zero entrance price. By adding the quantity demanded across all zones at entrance price p, we have the aggregate demand for the site:

$$v(p)=\sum_{i=1}^{m} v_i(p)=\sum_{i=1}^{m} P_i h(c_i+p, d_i, y_i). \tag{9}$$

Consumers' surplus, as measured with the aggregate demand curve, is

$$B=\int_0^{p^*} v(p)\,\mathrm{d}p, \tag{10}$$

where p^* is the reservation price. Expression (10) is equivalent to expression (8) providing p^* is asymptotic and hence the same for all distance zones or p_i^* is substituted for p^* for different zones. Graphically, (8) and (10) say that we may add up the area under the aggregate demand curve or add up areas under individual demand curves.

Figures 15.2 and 15.3 show per capita and aggregate demand curves. In Figure 15.2, the per capita demand for the site is drawn for a particular distance zone. To the extent that arguments such as distance (d_i) and income (y_i), as well as other arguments to be considered later, are appropriate and included in the per capita demand curve, each zone will have its per capita demand curve located differently. Expression (8) multiplies areas such as $p_i^* c_i A$ by zonal population and adds across zones to compute benefits. In Figure 15.3, the aggregate demand curve, $v(p)$ in (9), is drawn. The aggregate demand curve must slope downward but can have a variety of shapes because it depends upon the distribution of population. According to expression (10), site benefits may be computed as the area $Op^*v(0)$. These two benefit measures are identical.

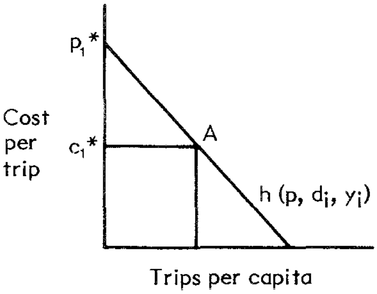

Figure 15.2. Per capita demand function.

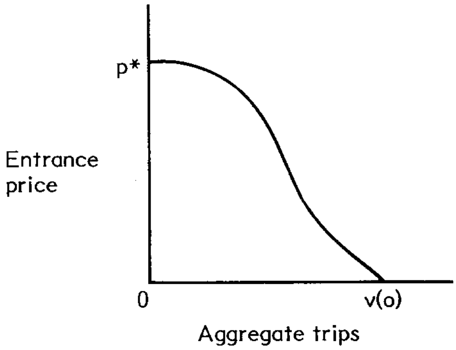

Figure 15.3. Aggregate demand function.

The mechanics of the travel-cost approach are simple and accepted in economics. The value of a good to consumers is approximated by the area under the aggregate demand curve above the prevailing price line. What is novel about the travel-cost approach is using the variation in the cost of reaching the site from different distances. Consequently, it is natural that the bulk of other issues arising in the travel-cost approach deal with the estimation of the travel-cost demand function. Empirical issues partly reflect the uses made of travel-cost models, primarily calculation of net benefits. Hence, much effort has been devoted to reconciling differences between models implied by utility-maximizing consumers with the need to complete analyses with imperfect data. The following subsections deal with some of the more important empirical problems which researchers have confronted.

3.1.1. Sources of data

Because the travel-cost method is an approach for estimating the demand for a specific site, it must have survey data which gives user destinations. Ignoring this requirement has led to some abuse of the technique. The nature of the difficulty can be explained by considering two common types of surveys: (1) the site-specific survey and (2) the population-specific survey. The travel-cost method is basically designed to be implemented with site-specific survey data.

Site-specific surveys are designed to gather visitation data at specific sites. The visitation rate may be measured in trips, day visits, or other systematic measures of service flow to users. Site-specific surveys can vary in form from on-site interviews to mail questionnaires which gather information about a specific site. The distinguishing characteristic of site-specific surveys is that they give the critical information for estimation of the travel cost: the dependent variable is computed from individual responses about visits to the particular site. Information about costs can be derived from the respondent's residence, miles traveled, or actual figures on the cost of a representative trip. The measurement of trips and costs per trip will be addressed below.

The population-specific survey is designed to gather aggregate measures of participation, usually in a variety of different recreational activities. These surveys tend to be national in scope, and gather information on nonparticipants. Activities covered in such surveys tend to occur on many different sites over many different geographical areas. Population-specific surveys, by their nature, involve sampling the population at large, rather than the population of users.

The population-specific survey measures attendance at many different sites in many different areas. Hence, these surveys are not useful for estimating site demand functions. Despite this rather obvious observation, there have been a number of efforts attempting to estimate travel-cost demand functions from population-specific surveys. The issue is essentially one of aggregation, where the aggregation is over demand functions for individual sites. However, the sum of demands is not a valid index of aggregate quantiity, nor is expenditures on all trips divided by the sum of all trips a useful index of prices. Furthermore, even if one could devise a valid aggregation scheme, it is not clear what purpose it would serve, because it would not show how welfare would respond to changes in attributes of specific sites.

3.1.2. *Utility maximization: Implications for measurement and specification*

The goal of the travel cost approach is to use site-demand functions to infer the consumer's surplus for access to the site. Because measurement of surplus is intimately related to maximizing behavior, it seems appropriate to turn to the model of utility maximization to aid in specifying travel costs demand functions. First we specify a simple model of consumer behavior, based on an elementary household production approach. Then we show how this model gives insight into the choice of measurement of the service flow and the cost per unit service flow, with special attention devoted to the cost of time.

Suppose that a representative consumer maximizes a quasi-concave utility function subject to a budget and a time constraint:

$$\max_{x,z} \{ u(x,z) \mid y = cx + pz, T = h + x(t_1 + t_2) \},$$

where

x = number of trips to a given site,
z = Hicksian commodity bundle of other goods,
h = time spent working,
t_2 = time spent on site per trip,

t_1 = travel time per trip,
T_0 = total time available,
y^0 = exogenous income,
$y = y^0 + wh$ = money income,
c = out-of-pocket costs per trip,
p = price of the Hicksian bundle, and
w = wage rate.

This elementary approach assumes that individuals can trade between work and leisure at a constant wage rate. Thus, $x(t_1 + t_2)$ and h are measured in the same units. For example, if h is hours worked, then t_1 and t_2 must be hours spent on site and hours traveling per trip, respectively. For the time, assume that there are no wedges created by income or other taxes. By combining the income and budget constraint, the problem becomes

$$\max_{x,z} u(x, z) + \left[y^* - c^*x - pz \right], \tag{11}$$

where $c^* = w(t_1 + t_2) + c$, the full cost, and $y^* = y^0 + wT$, full income. The first-order conditions for this problem include

$$\partial u/\partial x = \lambda c^* = \lambda \left(c + w(t_1 + t_2) \right), \tag{12}$$

and the relevant general demand function is $x = f(c^*, p, y^*)$. In a cross-section framework, p is the same for all observations. Hence the equation of concern is

$$x = f(c^*, y^*). \tag{13}$$

Several important estimation issues are revealed in this framework.

3.1.2.1. Measuring the variables. In applying the travel-cost method, one of the first decisions in survey work concerns the units of measurement of x. Trips should be of constant length. If t_2, the measure of time spent on site, is a single day, then trips per user and days per user as a dependent variable in (13) will yield the same result. However, the essential result of the choice process outlined here is that the full-cost argument described in (13) is the cost per trip, not cost per day. When trip lengths vary, the assumptions of the model are violated. The length of the trip (t_2) becomes a choice variable, and also needs an explanatory equation.

In practice, the difficulty caused by having observations on trips which vary in length can be handled by estimating separate demand functions for trips of different length. For example, we might have, for a remote resource, one-day demand functions, two-day demand functions, one-week demand functions, etc. This approach is explored in Brown and Mendelsohn (1984). The alternative is to make the length of the trip endogenous, so that two choices are made: (1) how many trips to take, (2) how long per trip to spend [see, for example, Smith, Desvousges, and McGivney (1983)].

The calculation of costs for the travel-cost method creates substantial difficulty. There are two problems: (1) the out-of-pocket costs and (2) the opportunity cost of time. The simplest travel cost models compute the out-of-pocket costs by multiplying round-trip mileage by cost per mile. The only source of trouble here is determining what cost per mile figure to use. For example, Bishop and Heberlein (1980) use 7.764 cents, 16.044 cents, or 15.684 cents, depending on whether the vehicle was compact, standard, or recreational (p. 39). Cicchetti, Fisher and Smith (1976) use two measures: (1) 1.7 cents per person per vehicle mile, based on 5.5 cents per vehicle mile and an average of 3.2 persons per vehicle; (2) 4.4 cents per mile (p. 1272). A summary of the costs per mile is given in Dwyer, Kelley and Bowes (1977) where, for 27 travel-cost models, the cost per mile of travel varies from 1.5 cents to 10 cents (table 1, pp. 45–51).

The impact on estimates of consumer's surplus of using the incorrect cost of travel, as well as other systematic impacts in the travel-cost method, can be determined relatively easily. Let the demand curve be

$$x = \beta_0 + \beta_1 c + \text{error}, \tag{14}$$

where β_i are the unknown parameters to be estimated. Using this demand curve, we calculate consumer's surplus to be

$$b = -x^2/2\beta_1.$$

When the actual cost is some proportion θ of the true cost, the expected value of the coefficient estimate is β_1/θ and the estimated consumer's surplus is θb. Thus, underestimating travel costs by, say, 25 percent also reduces the consumer's surplus estimate by 25 percent.

Evidence on the travel cost per mile is easily gathered, and it is clear that any travel-cost method must measure this component of the costs. Practices differ on other types of costs. Obviously, expenditures on parking entrance fees, tolls, and other expenditures necessary to reach the site should be included in the cost per trip. But what about expenditures on food, beverages, and other items which might be termed discretionary? Deciding what is discretionary is not a simple matter in economics. Some expenditures on goods are clearly part of the costs of traveling 300 miles. Likewise, for two-day trips, expenditures on lodging comprise a reasonable part of the costs per trip.

Suppose that some irrelevant expenditures are included in the cost per trip, c. We can examine the implications via a simple error-in-variables model. Suppose the measured and true cost per trip are related by

$$c' = c + c^0,$$

when c' is the measured, c the true, and c^0 the irrelevant cost per trip. When expression (14) is estimated, with c' substituted for c, the expected value of the estimator of β_1 may not be well defined but the probability limit of the estimator is given by

$$\operatorname{plim}\hat{\beta}_1 = \beta_1 \frac{\operatorname{plim}[m(c,c) + m(c,c^0)]}{\operatorname{plim}[m(c,c) + 2m(c,c^0) + m(c^0,c^0)]}, \tag{15}$$

where $m(z_1, z_2)$ is the covariance of z_1 and z_2. This result assumes that the error in (14) is independent of c^0 and c. From (15), we see that if the same irrelevant expenditure is added to each cost-per-trip figure, c^0 will be constant, $m(c^0,c^0) = m(c^0,c) = 0$, and $\operatorname{plim}\hat{\beta}_1 = \beta_1$; the estimate will be consistent. It will also be unbiased. When a random but irrelevant expenditure is added to each per-trip cost, then $m(c^0,c^0) > 0$, $m(c^0,c^0) = 0$, and

$$\operatorname{plim}\hat{\beta}_1 = \theta\beta_1, \tag{16}$$

where $\theta = \operatorname{plim} m(c,c)/\operatorname{plim}[m(c,c) + m(c^0,c^0)] < 1$. Since $\beta_1 < 0$, this measurement error makes $\hat{\beta}_1$ closer to zero than its true value, and the estimated consumer's surplus will be higher than its true value. When the irrelevant cost component is correlated with the true per-trip cost, the bias will depend on the probability limits of moments as in (15).

3.1.2.2. The cost of time. Measuring the opportunity cost of time is one of the fundamental research issues. It has been recognized as important from the inception of the method. For example, Clawson, in discussing the costs of visiting Yosemite, states, "The time required to visit Yosemite might well be a major cost to many potential visitors" (p. 21). Because the cost of time can be a major component of the total cost of visiting a site, incorrectly measuring it or excluding it can have a major impact on the measurement of consumer's surplus. This, too, has long been recognized. In discussing the impact of omitting the cost of travel time, Knetsch (1963) argues that "the method underestimates the number of visits that would occur with a postulated increase in costs, and consequently the demand derived in this manner is to the left of the actual curve.... The value imputed to the resource is therefore also underestimated" (pp. 395–396).

Thus, the problem is fairly well known and its consequences apparent. Yet, a satisfactory resolution of the various aspects of the problem remains to be discovered. There are two aspects to the cost of time issue. One is how to measure the cost of time. A second is how much time should be measured as costly, i.e. travel time, on-site time, or both. A related aspect is the issue of whether, given that on-site time is to be calculated as part of the cost, its opportunity cost is the same as that of travel time.

To capture the essence of the cost of time issue, suppose that the demand function (13) can be specified as linear, and that income is known to be irrelevant. An alternative to assuming income irrelevant would be to assume income orthogonal to distance from the site. Under these assumptions, a stochastic specification of (13) would be

$$x_i = \beta_0 + \beta_1(c_i + w_i(t_1^i + t_2)) + \text{error}, \tag{17}$$

when the index i refers to individual demands. Note that t_1^i, the travel time, by necessity, varies among individuals, but that t_2, time on site, does not vary. This difference immediately distinguishes travel time and on-site time, for, if the opportunity cost of time is constant across individuals ($w_i = \bar{w}$), we can write (17) as

$$x_i = \beta_0 + \beta_1 \bar{w} t_2 + \beta_1(c_i + wt_1^i) + \text{error}.$$

Hence, when time on site is constant and the wage rate is constant, the cost of time on site is constant across individuals, and becomes part of the constant term in a linear regression equation.

Suppose that instead of the true model of (17), we estimate

$$x_i = \hat{\beta}_0 + \hat{\beta}_1 c_i + \text{residual}. \tag{18}$$

The expected value of $\hat{\beta}_1$ is

$$\mathrm{E}\hat{\beta}_1 = \beta_1[m(c,c) + m(c,wt_1) + t_2 m(c,w)]/m(c,c), \tag{19}$$

assuming no correlation between the error in (17) and the regressors. The bias depends on the signs and size of $m(c,wt_1)$ and $m(c,w)$. If we make the traditional assumption that w_i is constant, or at least uncorrelated with costs and time from the site, (19) becomes

$$\mathrm{E}\hat{\beta}_1 = \beta_1[m(c,c) + \bar{w}m(c,t_1)]/m(c,c). \tag{20}$$

The $m(c,t_1)$ term is the correlation between out-of-pocket costs per trip and travel time per trip, which is surely positive, yielding $\hat{\beta}_1$ biased upward in absolute value. Using a biased estimate of β_1 leads one to infer that trips respond more to increases in costs than they actually do, implying that users can be priced out of the site more easily than they actually can, and an underestimate of the value of the site.

A comment about the specification of (17), where the income variable is left out, is in order. It is a common finding for site demand functions that the income effect is insignificant. For example, in Desvousges, Smith, and McGivney, results for 43 site demand functions show that in only two cases does the ratio of coefficient to standard error for the income variable exceed 1.9. There are two reasons for this empirical finding. First, because demand functions are site-specific, higher income may simply lead to the selection of other sites. Second, when the cost of time is not accounted for in the demand function, the income variable measures two offsetting effects: the expansion of the budget constraint and the substitution effect caused by higher wages which generally occasion higher income. Hence, it is biased toward zero.

The assumption that leisure can be substituted for work at a constant wage rate is pivotal in the model as formulated. Under this assumption, the wage rate is a good measure of the value of time. However, even a cursory look at behavior and institutions forces the admission that this model is inaccurate. First, the income tax creates a substantial difference between before and after tax wage rates, making the opportunity cost of time less than the wage rate. Second, people who get utility (or disutility) from work will need to be compensated more (or less) at the margin than the after-tax wage rate. Cesario (1976) uses the notion that people get disutility to argue that the marginal value of time is approximately equal to one-third of the wage rate. McConnell and Strand (1981) argue that the *a priori* evidence on the value of time is inconclusive, especially when the wage rate is often computed by dividing reported income by 2000, the number of hours in a typical work year. They suggest that expression (17) be approximated by (with observation indices suppressed)

$$x = \beta_0 + \beta_{11}c + \beta_{12}t_1 w + \text{error},$$

which can be rewritten as

$$x = \beta_0 + \beta_{11}[c + kt_2 w] + \text{error}, \tag{21}$$

where $k = \beta_{12}/\beta_{11}$, the proportion of the wage rate which is the marginal value of time. McConnell and Strand estimate k to be about 60 percent, based on a disaggregated model of salt water anglers. Smith, Desvousges and McGivney (1983), however, show that when computed over many sites, k ranges from -9 to 80.

The opportunity cost of time is traditionally conceived as a parameter, but it need not be so. When people give up work at the constant marginal wage rate, the cost of time is a parameter. However, when the opportunity cost of time is created by other nonmarket alternatives, the opportunity cost of time is no longer a parameter. Let us demonstrate this by revising the choice model so that an individual is given his income, but uses his leisure time to choose recreational sites. Let the consumer maximize

$$u(x_1,\dots,x_S) + \lambda\left(y - \sum_{j=1}^{S} c_j x_j\right) + \mu\left[T - \sum_{j=1}^{S}(t_{1j} + t_{2j})x_j\right], \tag{22}$$

where x_j, t_{1j}, t_{2j}, and c_j refers to trips, travel time, on-site time, and travel costs to the jth site, and S is the number of sites. T is the total time available for recreation, and μ is the marginal utility of additional recreation time. The first order conditions for this problem can be written

$$\partial u/\partial x_j = \lambda\big(c_j + m_T(t_{1j} + t_{2j})\big), \qquad j = 1, S, \tag{23}$$

where $m_T \equiv \mu/\lambda$, the marginal value of time. This expression has a superficial similarity to (12), when w is substituted for m_T. However, in (23), the marginal value of time is not a parameter, but a ratio of endogenous variables, μ/λ. In that case, individual site demand functions, which are the solutions of (23) and the time and budget constraints, can be written

$$x_j = f_j(c, t_1, t_2, y, T), \tag{24}$$

where c is a vector of out-of-pocket costs, t_2 is the vector of on-site times, and t_1 is a vector of travel time. In (24), the marginal value of time is no longer a parameter. The parameters are the costs and time requirements for each site. Although the value of time is crucial to decisions in the multiple-site case, it is an outcome of decisions. Measurement and collinearity make a specification such as (24) little more than a formalism. The time and budget constraints can provide guidelines for research, but they cannot provide explicit answers. A further difficulty with the model as formulated in (22) is that elementary comparative statics afforded by a single constraint may not hold when two constraints are involved.

Recent work by Smith, Desvousges and McGivney (1983) extends empirical research on the value of time in several directions. First, they develop the choice model in a framework in which the marginal value of time is not equal to but, in part, determined by the wage rate. Second, they use a hedonic price equation to predict the wage rate, a practice preferable to the approach of dividing reported family income by the number of hours available for work per year. Third, they develop models in which on-site time is part of the choice process. Further work by Bockstael, Strand and Hanemann (1984) develops a model for the most realistic case, where some recreationists have wage-earning opportunities at the margin, while others do not.

3.1.2.3. The role of substitutes. As with other economic goods, costs of substitutes play a crucial role in determining the demand for outdoor recreation. And in the estimation of outdoor recreation demand functions, costs of substitutes may be highly correlated with the costs of visiting a site, especially when substitutes are not distributed uniformly in a region.

The omission of relevant costs of substitutes may bias estimates of coefficients on the included variable. Suppose, for simplicity, that the true model is

$$x = \beta_0 + \beta_1 c_1 + \beta_2 c_2 + \text{error}, \tag{25}$$

and the estimated model is

$$x = \hat{\beta}_0 + \hat{\beta}_1 c_1 + \text{residual}, \tag{26}$$

when c_1 is the on-site price, c_2 is the substitute site price. Assuming that the regressors are independent of the error in (25), we compute the expected value of $\hat{\beta}_1$:

$$\mathrm{E}\hat{\beta}_1 = \beta_1 + \beta_2 m(c_1, c_2)/m(c_1, c_1). \tag{27}$$

If the sites are substitutes, $\beta_2 > 0$. If the sites are located in opposite directions from each other, $m(c_1, c_2) < 0$. In the case where $\beta_2 m(c_2, c_1) < 0$, $\hat{\beta}_1$ will be biased away from zero, leading to the inference that trips per user respond more rapidly to an increase in site entry costs than, in fact, they do, and hence to an underestimate of consumer surplus.

The proper handling of recreational substitutes depends on the nature and uniqueness of the resource, as well as the issue being addressed by the analyst. Obviously, there are myriad substitutes for any kind of leisure activity, but many can be ignored because their prices do not vary systematically across individuals. For example, going to the movies can be a substitute for going to a lake going fishing. But while the cost of going fishing will vary systematically with distance from the site, the cost of going to the movie may not. Hence, while changes in the price of movies could influence the level of fishing activity, this price can be safely omitted from the demand for the fishing site equation.

The relative uniqueness of a resource is an important determinant of the role of substitutes. For example, the Grand Canyon has many substitutes, but few which vary systematically with distance from the site. Campgrounds in Georgia, New York, and Illinois all provide substitutes for the Grand Canyon. Yet, the prices of these substitutes do not vary systematically with distance from the Grand Canyon. Thus, one might reasonably argue that there is no serious bias when a demand function for the Grand Canyon is estimated using observations on persons throughout the United States.

Studies which have successfully estimated cross price effects provide insight into the circumstances when it is useful to include the price of substitutes [e.g. Burt and Brewer (1971), Cesario and Knetsch (1976), and Cicchetti, Fisher and Smith (1976)]. Burt and Brewer develop a multiple site demand model to evaluate a proposed new site. They estimate the demand for six types of water-based recreation. One type of site was much like the Army Corps of Engineers lake being proposed. A system of six demand functions of the following sort is estimated:

$$x_{ij} = \alpha_j + \sum_{k=1}^{5} \beta_{jk} p_{ik} + \sigma_j y_i + \varepsilon_{ij}, \tag{28}$$

where

x_{ij} = trips by ith household to jth site group,
p_{ik} = price to ith household of visiting kth site,
y_i = income of ith household,
α, β, σ = parameter vectors to be estimated, and
ε_{ij} = random error.

The parameters of this system were computed using a constrained estimation technique. The symmetry constraints, that $\hat{\beta}_{ij} = \hat{\beta}_{ji}$, were imposed to ensure a unique money measure of welfare. These constraints, $\partial x_i/\partial p_j = \partial x_j/\partial p_i$, imply that the subutility function defined on waterbased recreation services is homothetic. With the estimated coefficients, the benefits of the new site are computed by assuming that the new site is identical to some existing sites. Since both the new and some of the existing sites are Corps of Engineers lakes, this is a reasonable assumption. Because the sites are almost identical, the sole implication of introducing a new site is a lower price for this type of site. Benefits are computed as the change in the area under the demand curve for the perfect substitutes for the new site.

The Cicchetti, Fisher and Smith study, similar in spirit to the Burt and Brewer work, evaluates a proposed new skiing resort in California. The benefits of the new resort are determined by computing the increased area under the demand curve for perfect substitutes. Since their observations were per capita use by county in California, some counties gained by having the new site closer, while other counties did not.

Estimation of systems of demand functions has two impacts on benefits. First, the addition of the cost of visiting substitute sites to each demand function influences the estimate of the coefficients of the own price variable. Second, by recognizing that the introduction of a new site which is a perfect substitute for existing sites will provide increased benefits to some observations, but not to others, the multiple site approach tends to reduce estimates of benefits.

3.1.3. Aggregation

Little attention has been given to the measurement of the dependent variable. Here we address the issue of whether individual observations should be used by themselves or aggregated to per capita observations. This issue, the aggregation problem, has substantial impact on coefficient estimates. The initial focus on aggregation comes from Brown and Nawas (1973).

The aggregation issue, though treated separately here, arose because of other difficulties in the travel cost model. The original travel cost model used trips per capita by distance zone as observations per capita. Early attempts to deal with the value of time and the availability of substitutes involved using distance as well as cost per trip as an argument in the demand function. The obvious high collinearity between distance and cost per trip make this approach particularly unworkable when the number of observations is small, as it is likely to be when zones are used for observations.

Brown and Nawas (1973) show that using individual observations within a distance zone rather than grouping the observations increases the efficiency of parameter estimates, reduces multi-collinearity, and permits the estimation of parameters on distance and cost in the same equation. Brown and Nawas present two equations:

$$\frac{\sum_i^{n_j} x_{ij}}{n_j} = 2.4 - \underset{(1.25)}{0.0087}\frac{\sum_i^{n_j} c_{ij}}{n_j} - \underset{(3.75)}{0.0079}\, d_j + \text{residual} \tag{29}$$

and

$$x_{ij} = 2.4 - \underset{(4.6)}{0.009}\, c_{ij} - \underset{(6.6)}{0.0069}\, d_j + \text{residual}, \tag{30}$$

where x_{ij} is the number of big game hunting trips and c_{ij} is the cost per big game hunting trip for the ith household in the jth distance zone to the hunting site in the Cascade Mountains of Oregon. The variable n_j is the number of sample hunting households in the jth zone and d_j is the distance from the jth zone to the hunting site. In eq. (29) we have the travel-cost method as originally conceived. It uses average household trips from 31 distance zones, and has an uncorrected R^2 of 0.6. The t-statistic (in parenthesis) under the null hypothesis of no association on the average cost variable is low, and the variable is not significantly different from zero at the 90 percent level. Eq. (30), estimated from 248 individual observations on hunters, explains only about 30 percent of the variance of trips per user. However, both the cost variable and the distance variable are significantly different from zero at a high level of confidence.

Though the specifications of expression (30) are not generally used, the procedure of using individual observations rather than zonal averages has become a common practice. [See, for example, Gum and Martin (1975), Smith, Desvousges and McGivney (1983), and McConnell and Strand (1981).] The most serious difficulty in using (30) rather than (29) arises from the errors-in-variables problem. From expressions (15) and (16), we see that the degree of bias in the OLS estimate of the cost coefficient is proportional to the variance in the error of costs per trip. Following Brown et al. (1983), this bias can be reduced by using a zonal average for costs in the individual demand model. For example, if we think that users from a particular zone have the same cost, but each individual cost is subject to random error, then instead of the specification in expression (30) we could use

$$x_{ij} = \beta_0 + \beta_1 \sum_i c_{ij}/n_j + \text{error}. \tag{31}$$

Brown et al. show that even when the number of users sampled per zone (n_j) is relatively small, there are important gains from using an expression such as (31) in which the cost per trip error is reduced by averaging. In the case considered by Brown et al., there are at least five users sampled per zone, implying a reduction of 80 percent in the variance of the random component of trip costs.

There are, as always, uncertain tradeoffs to make in the decision to aggregate, even if it only involves aggregating and averaging costs per trip as in expression

(31). The difficulty that arises in (31) is that it assumes that the cost per individual per trip is the same for all individuals from a given distance zone. However, the travel costs, reasonably assumed to be constant for individuals within a distance zone, are only one component of the cost per trip. It is quite plausible that the costs of time vary across individuals within a distance zone. If so, the averaging of costs will involve a loss of information. Since we have not yet discovered an acceptable method for measuring the opportunity cost of time, however, the approach of (31) seems a reasonable working solution.

In contemplating the aggregation problem, we might be mindful of the advice given by Hicks (1956) on statistical demand functions: "To assume that the representative consumer acts like an ideal consumer is a hypothesis worth testing; to assume that an actual person, the Mr. Brown or Mr. Jones who lives round the corner, does, in fact act in such a way does not deserve a moment's consideration" (p. 55).

3.1.4. Defining distance zones

When the researcher goes to the field to gather data, or is presented data from which a model must be estimated, he must define distance zones in such a way that sample observations may be clearly defined. Although the travel cost model was originally conceived in terms of concentric zones around a site, it is often now practical to use political jurisdictions; for example, counties are often used as distance zones. With counties as observations, the dependent variable is some measure of per capita use, and independent variables can be drawn from independent sources as well as from the sampled population.

Several different issues must be dealt with once distance zones are defined. An issue for which there is guidance but no clear solution concerns how far to extend the region of analysis, i.e. how many distance zones should be included. A second issue deals with the distribution of the error term when per capita data are used.

The problem of determining how many distance zones to use in the estimation of a travel-cost model arises because of the likelihood that users further from the site may have different behavioral parameters. The travel cost model is based, in part, on the assumption that trips are made with the sole purpose of visiting the recreational site. However, "As we expand the set of origin zones to include progressively more distant units the assumption that each trip is a single purpose excursion becomes more untenable" [Smith and Kopp (1980, p. 65)]. For some national parks, Haspel and Johnston (1982) show that the multiple purpose trip is an important issue. The Smith–Kopp approach to dealing with the spatial limits of the travel cost method stems from defining the basic model as

$$x_j = \beta_j w_j + \text{error}, \qquad j = 1, m,$$

where x_j is trips per capita from zone j, w_j is the vector of exogenous variables, including costs per trip, from zone j, β_j is the vector of parameters to be estimated for zone j, and m is the number of zones. Smith and Kopp adapt a statistic based on successive computation of orthogonal residuals, beginning at the site, to show when estimated parameters make a significant change. When a significant change occurs, the remaining zones should be dropped. The test does not, of course, allow the estimation of different parameter vectors. For the case analyzed by Smith and Kopp, the consumer's surplus per trip is reduced from \$14.80 to \$5.28 because of the deletion of the approximately one-third most distant zones. The Smith–Kopp approach is a useful tool when the combined on-site and travel time exceeds one day, so that multiple purposes for the most distant zones are possible. The technique is also useful for data sets which have zero trip levels for some of the more distant zones.

When per capita data are used to estimate the travel cost model, it is implausible for a variety of reasons to assume that the variance of the error will be the same for all distance zones. Consider the travel cost model

$$x_i = \beta w_i + \varepsilon_i, \tag{32}$$

where x_i is trips per capita per distance zone and $\varepsilon_i =$ zonal random error. This distribution of ε_i depends on the data source. When (32) is estimated from data which are aggregated across the population of the ith zone, and each individual has the same random error distribution for trip demand, then $V(\varepsilon_i) = \sigma^2/n_i$, when σ^2 is the individual variance and n_i is population of the ith zone. Hence the model is heteroscedastic. Bowes and Loomis (1980) argue that generalized least squares can be used to correct for the heteroscedastic error in (32) by multiplying data for the ith distance zone by $\sqrt{n_i}$. As McConnell and Bockstael (1984) show, however, this simple expedient works only when the participation rate is constant over distance zones. And any plausible derivation of the participation rate would show that it declines with distance.

3.1.5. Functional form

The problem of heteroscedasticity in the use of zonal observations can be detected in a rough and ready way by plotting the OLS residuals against variables suspected of creating the heteroscedasticity. But this sort of test is also a useful way of getting a feel for whether the estimated functional form is appropriate for the given set of data. The similar symptoms have led several researchers to test for functional form as well as heteroscedasticity [Vaughan, Russell and Hazilla (1982), Strong (1983)].

Initial research with the functional form stemmed from concern over the role of income and the distribution of income. Seckler (1966) asked whether measuring the value of recreation sites using travel cost models was a legitimate use of economics. Seckler's concern was that consumer's surplus reflected income, and hence for supposedly superior goods such as recreation, this criterion would allocate goods toward upper income groups. This concern spurred an interest in determining how income influences the demand for trips, leading to the modest finding that functions linear and additive in price and income cannot be derived from utility maximization [McConnell (1975)].

Current research is concerned with the functional form for two related reasons. First, differences in functional form can make a substantial difference in the consumer's surplus. For example, in Ziemer, Musser and Hill (1980), the estimated consumer's surplus for a linear demand source is about three times that of a semi-log function. Second, different functional forms imply different relationships between Hicksian and Marshallian surpluses, and, in fact, some functional forms create difficulties with the bounds derived by Willig (1976). For example, see Bockstael and McConnell (1979) and Hanemann (1980) where the impact of several popular functional forms on the divergence between Hicksian and Marshallian surplus is debated. Thus, the source of concern over the functional form has shifted from the issue of equity – that income differences make willingness to pay an inappropriate criterion for evaluating recreational sites, to the issue of efficiency – that the functional form and its implied individual preference function may not give the correct measure of what an individual will pay or accept for compensation.

Because there is little theoretical guidance in choosing functional forms, researchers have utilized statistical techniques. Results from statistical tests are somewhat uniform and tend to corroborate the initial findings of Smith (1975). In this work, Smith uses Box–Cox techniques on individual observations to select approximately the form

$$\ln x = \beta \cdot w + \text{error}.$$

This functional form is also selected by Ziemer, Musser and Hill (1980) using individual observations, Strong (1983) on zonal observations, and Vaughan, Russell and Hazilla (1982) using zonal observations.

Both the Vaughan et al. work and the Strong work on functional form show that tests for heteroscedasticity of the sort suggested by Bowes and Loomis are strongly dependent on the maintained hypothesis concerning functional form. In an analysis of the original Bowes and Loomis data, Vaughan et al. reject linear homoscedastic and heteroscedastic models in favor of models which are approximately semi-log. However, given a wider range of maintained hypotheses about the functional form, they are unable to choose between a homoscedastic and a heteroscedastic error structure.

In conclusion, choice of functional form is important. As well as leading to computation of consumer's surplus which accords with behavior, the correct functional form apparently alleviates the problem of heteroscedasticity. And while there is no particular reason to believe that a functional form which holds for one recreational site will work best for another site, a researcher who was forced to choose from the literature rather than be guided by theoretical or statistical considerations would find that the bulk of the evidence supports a semi-log form.

This chronicle of the estimation issues for the travel cost method has dealt with the major issues. In its current state, the travel cost method for estimating recreational benefits reflects two distinct forces. One force is the apparently growing demand by the public sector for methods to evaluate natural resources on the basis of their benefits to individuals. This force gives rise to a large but indiscriminate demand for techniques which will yield measures of benefits. Another force is the intellectual debate which asks not whether tools are pragmatic or how resources are allocated, but under what circumstances results stand the test of reason. The tension between the two sources is natural, and in the case of the travel cost method has created a valuable tool for resource allocation which is still evolving.

3.2. Contingent valuation

The contingent, or hypothetical, evaluation approach to measuring economic surplus derived from access to a recreation site attempts to measure directly the surplus rather than compute it from a demand curve. The contingent valuation approach is based on the belief that individuals are capable of responding to questions so as to reveal their preferences for public goods. The literature on contingent valuation has become voluminous and diverse for at least two reasons. First, this method has broad applicability and is used for many different kinds of resources, for example, hazardous waste, visibility, water pollution. It can be used for resources which have not yet been developed, as well as public goods which are not complementary to any private activity. Second, there is a lingering controversy stemming from the hypothetical nature of the approach: contingent valuation seeks to use responses to hypothetical questions to measure the benefits for a public good. Because the contingent valuation approach is much broader than the economics of outdoor recreation, it will be treated here only to the extent that it deals with outdoor recreation (additional discussion of the contingent valuation approach can be found in Chapter 6 of this Handbook).

3.2.1. The basics of contingent valuation

The application of contingent valuation methods requires recreationists to state their willingness to pay or willingness to be compensated for access to a particular recreation site. To carry out the approach, the researcher must design a questionnaire which will induce the recreationist to reveal his willingness to pay (or be compensated) and devise a sampling procedure which will confront the recreationist with the questionnaire. Sampling procedures have tended to be on-site surveys of recreationists, but home and phone interviews and mail surveys have also been used.

The substance of contingent valuation is that respondents are required to give a hypothetical answer to a hypothetical question. The answer, in terms of consumer's surplus, is a number which, by its nature, cannot be observed. Hence, especial care with the design of the survey instrument seems warranted. Two important aspects of the instrument are: (1) the way the respondent's answer is determined and (2) the vehicle through which the respondent conceives of his payment being made.

There are myriad ways in which questions can be put to the respondent. Principal approaches are the iterative bidding game [Knetsch and Davis (1966), Randall, Ives and Eastman (1974)], the open-ended question [Hammack and Brown (1974)], and a check list [Desvousges, Smith and McGivney (1983)]. In the iterative bidding game, the respondent is presented with a benefit figure (maximum willingness to pay or minimum willingness to accept compensation) which is increased or reduced until the individual reaches his equilibrium. In the open-ended question, the individual is simply asked to give his maximum willingness to pay or minimum to accept compensation. With a check list, the respondent simply checks the correct benefit figure from a list.

There are, as well, different media through which an individual may conceive of making his payment. These media, known as vehicles, are quite diverse. Examples are sales tax, income tax, entrance fees, changes in income, trip costs, utility price increases, and increases in the costs and prices of all goods and services. For example, in the 1980 National Survey of Hunting, Fishing and Wildlife-Related Recreation, an iterative bidding game was executed in interviews administered at the home of the respondent. The vehicle was trip costs, and the method of questioning required the interviewer to raise or lower the hypothetical cost per trip until the per unit cost which induced zero trips was discovered.

The original work using contingent valuation on outdoor recreation is a study by Davis (1963) on the Maine woods. The results of this study are summarized in Knetsch and Davis (1966). The approach is described as a "method of measuring recreation benefits... that through a properly constructed interview approach one can elicit from recreationists information concerning the maximum price they would pay in order to avoid being deprived of the use of a particular area for whatever purpose they may make of it" [Knetsch and Davis (1966, p. 131)]. The Knetsch and Davis analysis of the contingent valuation is well based in theory, and is an example worth examining. The interviews were administered in the woods of northern Maine and they "included a bidding game in which respon-

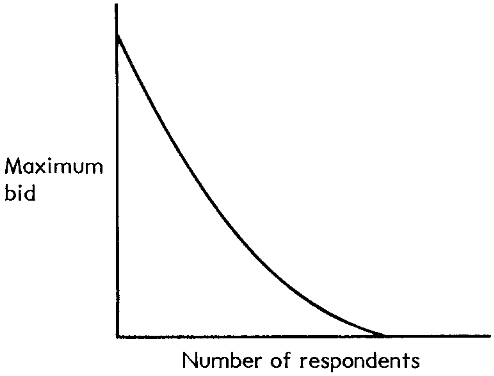

Figure 15.4. An aggregate demand function from contingent valuation.

dents could react to increased costs of visiting the area. Bids were systematically raised or lowered until the user switched his reaction from inclusion to exclusion" (Knetsch and Davis, p. 133). In addition to the bids, information on socioeconomic variables was also gathered during the interview. This information was used to estimate an equation which would predict the bid of individuals who were interviewed briefly but who did not participate in the iterative bidding game.

The survey results are used in an interesting way to derive the aggregate demand curve and hence consumers' surplus. Consider Figure 15.4. On the horizontal axis, measure the number of respondents. On the vertical axis, measure the maximum bid per respondent. Then any point on the schedule shows all the respondents who would not bid more than the point on the vertical axis, in effect the aggregate demand schedule. Using this aggregate curve, Knetsch and Davis estimate annual benefits of $71 461 for access to an area near Moosehead Lake, Maine. The fact that each user took only one trip to the site permitted the derivation of this aggregate schedule.

The Knetsch and Davis study is the first application of contingent valuation tools to outdoor recreation, but it is unique for other reasons. First, it also calculated benefits via the travel-cost method, allowing a comparison of the benefits by the two separate approaches. The benefits estimated by the travel-cost method were $69 450. Second, administering the bidding game to only a portion of the sample, and expanding the sample to the population by using a larger sample not including bids, suggests an efficient method of extrapolation. Third, Knetsch and Davis use the contingent valuation to derive explicitly the aggregate demand for the site, which, under certain assumptions, might have permitted the comparison of the predictive ability of the iterative bidding game with that of the travel-cost method.

3.2.2. Issues in implementing the contingent valuation method

Despite the substantial experience with contingent valuation by different researchers in different settings for different resources, there remains some skepticism about the validity of the results. This skepticism stems primarily from the belief that people do not know their economic surpluses for environmental assets. In contrast to market-based measures, such as the travel cost method, contingent valuation requires that users know approximately the worth of resources or changes in resources to them. Further difficulty is created by the fact that there is no true valuation with which the bid can be compared.

A considerable literature aimed at refuting or validating contingent valuation responses has developed. Three basic approaches to this problem have emerged. The first approach attempts to discover whether strategic motives, instrument design, or the nature of the vehicle lead to unreliable results. The second approach compares the results of contingent valuation with a measure of willingness to pay or accept compensation based on actual payments. A third approach compared benefits measured via the contingent valuation approach with benefits via the travel cost method. These approaches will be examined briefly.

3.2.2.1. Survey design. The basic issue is whether superficial variations in the questions can lead to significant variations in responses. The possibilities for survey-induced error have been neatly categorized as biases, including strategic, starting point, vehicle, and hypothetical. Strategic bias might arise if an individual thought that his responses would influence policy. For example, a high bid might increase the supply of environmental assets. Starting point bias occurs in an iterative bidding game where the final bid is conditioned on the starting bid, which is given to the respondent by the interview process. Vehicle bias occurs when answers vary with the vehicle. For example, it might be contended that using a tax as a vehicle would induce a response to high taxes, as well as evaluating the resource. Hypothetical bias is simply the difficulty of getting consistent and accurate responses to hypothetical questions.

The evidence amassed concerning these biases is mixed but improving. (See Desvousges, Smith, and McGivney for a summary.) But the problems addressed by the categories of bias pertain to survey design and not to the validity of contingent valuation approaches. Even if we know that some survey designs give more consistent approaches, we do not know which ones give the correct answers. Hence the emergence of other approaches to test contingent valuation responses.

3.2.2.2. Contingent valuation and cash exchanges. There is only one study which compares contingent valuation results with comparable cash payment for outdoor recreation: Bishop and Heberlein (1979, 1980). In that study, several different valuation approaches are used to value the right to hunt geese during the open season in the Horicon Reserve of Northern Wisconsin. A mail survey offered sportsmen checks of varying amounts in return for their hunting permit. Hunters were also asked direct questions about their willingness to pay and their willingness to accept compensation for their right to hunt. The results of the Bishop and Heberlein study show that the direct questioning form of contingent valuation yields answers for willingness to pay less than the cash transaction for willingness to accept compensation, which is, in turn, less than the stated willingness to accept compensation. The mean values of hypothetical willingness to pay, actual cash payment, and hypothetical willingness to sell were \$21, \$63, and \$101, respectively [Bishop and Heberlein (1979, p. 929)]. For relevant ranges of income elasticities, the willingness to pay an actual cash payment seem too far apart to be explained by Willig's results, suggesting that hypothetical willingness to pay is biased down.

The Bishop and Heberlein study has received mixed receptions. Some have used their results to argue the validity of contingent valuation approaches. Others have used the fact that actual willingness to accept compensation is about 60 percent of stated willingness to accept compensation to criticize the contingent valuation approach. As with any study of nonmarket benefits, there are fertile grounds for criticizing the Bishop–Heberlein study. It was a mail survey. It relied on results from an unusual market. There is no evidence that long-run equilibrium would produce the same results. This study nevertheless remains the only real test of the validity of contingent valuation approaches to benefit measurement.

3.2.2.3. Comparison of travel cost and contingent valuation answers. Relative to tests of actual cash payments vs. contingent valuation results, there are many studies which compare contingent valuations with the travel cost approach. For example, Bishop and Heberlein also estimate a travel cost model in their study, giving a benefit estimate of between \$11 and \$45 per permit, depending on the value of time. Sellar, Stoll, and Chavas (1983) find that the travel cost method and contingent valuation yield similar results when applied to Texas lakes. The general thrust of these comparisons of methods seems to suggest that there are not order of magnitude differences.

The finding that the travel cost method and contingent valuation approaches yield roughly similar measures of benefits should not be too reassuring, especially when one considers the many solutions to problems which arise in estimating travel cost models. Such comparisons only detect differences because the true value of benefits is, in general, not known.

In sum, the growth in the use of contingent valuation has resulted from the need for a valuation tool where observations on behavior cannot be simply gathered and from frustration with the myriad econometric problems of the travel cost method. Despite its popularity and the presumption of validity endorsed by widespread use, contingent valuation remains a tool for resource valuation which primarily reflects the demand by public sector for benefit evaluation. It has not passed the intellectual scrutiny which will help determine whether the method stands the test of reason. However, there is reason to believe that such a process is now beginning, and that eventually contingent valuation will be pragmatic, and at least for outdoor recreation resources, will yield testable results.

3.3. *The household production function*

It is intuitively appealing to construct a model of recreation behavior in which an individual buys certain private inputs, and combines them with his own time and publicly provided natural resources to produce outdoor recreation services. This model, the household production function, has been explored in Deyak and Smith (1978), Bockstael and McConnell (1981, 1983), and Brown, Charbonneau and Hay (1978), but has proved to be too general to yield readily available benefit estimates. It is nonetheless worth examining briefly because it can be seen that the travel cost approach is a special case of the household production function. A simple version of the household production process, in which an individual produces a recreation service flow at a single site, is presented.

Suppose that an individual buys a vector of market inputs at market prices, and combines them with a public resource (q) and his own time to produce recreational services, measured in days as x:

$$x = x(z, q).$$

The inputs z may be such items as miles driven, services of toll roads, and parking lots, and other private inputs which are part of the recreation package. They are purchased at market prices, denoted r. For simplicity, assume that time is purchased at its opportunity cost, the wage rate. Hence time is simply one dimension of z. The problem of choosing the level of service flow, x, can be conceived as a two-stage maximization problem. In the first stage, the recreationist minimizes the cost of producing a given level of services. The result is the cost function:

$$c(x, r, q) = \min_{z} \{ rz \mid x = x(z, q) \}. \tag{33}$$

In the second stage, the recreationist maximizes utility subject to the budget constraint $y = c(x, r, q) + p_h z_h$, when y is money income, z_h is a bundle of Hicksian goods, and p_h its prices. The problem

$$\max_{x, z_h} \{ u(x, z_h) \mid y = c(x, r, q) + p_h z_h \} \tag{34}$$

includes the first-order conditions:

$$\lambda^{-1} \partial u(x, z_h)/\partial x = \partial c(x, r, q)/\partial x, \tag{35}$$

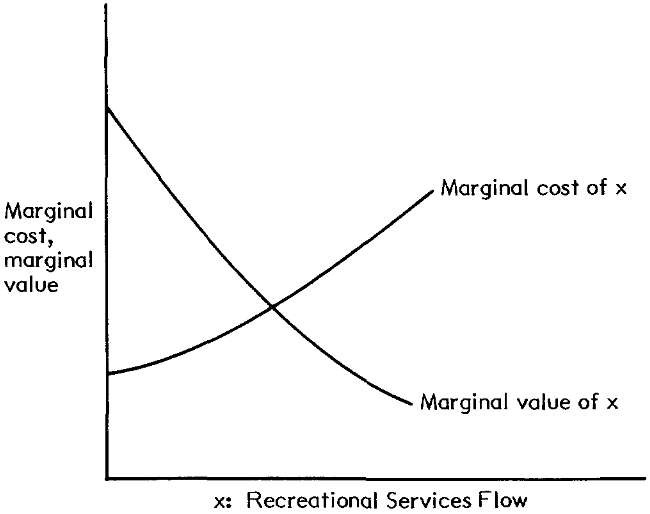

Figure 15.5. Marginal cost and value schedules for household produce services.

where λ is the Lagrangian multiplier associated with (34). In general, expression (35) implies a marginal cost and marginal value function for services as pictured in Figure 15.5. Within each household, there is an implicit market in which the equilibrium flow of recreational services is determined by the intersection of the marginal cost and marginal value functions. Pollak and Wachter (1975) show that, unless the production function has constant returns to scale and there is no joint production, the marginal cost of x will depend on x. When the marginal cost of x is variable, as is shown in Figure 15.5, simultaneous estimation of the marginal cost and marginal value function is necessary. Identification of the functions requires that some exogeneous variables be excluded from each function, a condition difficult to establish in practice. Estimation and benefit computation problems are covered in detail in Bockstael and McConnell (1981).

The simple travel cost model is a special case of the household production function. The only input to the production process is round trip transportation service to and from the site. Let the cost per trip in terms of transportation services be r_1. Then the implied cost function is

$$c(x, r_1) = xr_1.$$

The marginal cost is $\partial c/\partial x = r_1$, the round trip travel cost, is constant, and hence horizontal in Figure 15.5. Variations in r_1 identify the demand curve.

In conclusion, the household production function is an intellectually appealing way to think about the provision of recreational service. However, only in special cases are the estimation problems sufficiently worked out to make the tool an effective approach to measuring the benefits of outdoor recreation.

3.4. *Forecasting the demand for outdoor recreation*

The three previous subsections have surveyed methods used to estimate the economic value of outdoor recreation facilities. The first section gives evidence to show that federal and state government provide a significant portion of the resources available for outdoor recreation. Readers familiar with management and provision of resources by the public sector for outdoor recreation will not find it surprising that the impact of these values on policy decision is not always apparent. Thus, while it seems quite appropriate that economists should devote effort to measuring the economic values, there are other tasks in this area which can honorably engage economists' skills. An especially useful topic has been the forecasting of demand for outdoor recreation.

The need for forecasting the demand for outdoor recreation arises in the long-run planning of resources. For example, at the federal government level, acquisition of wildlife habitat requires some knowledge of aggregate use levels in the future. State governments rely on forecasts of various outdoor recreation activities to aid in planning the development of parks, beaches, and campgrounds, and to justify such developments. The need for forecasting at the state level is institutionalized by the requirement that each state develop a State Comprehensive Outdoor Recreation Plan in order to qualify for certain federal funds for the purchase of natural resources.

Forecasting models are less firmly embedded in neoclassical choice models than benefit evaluation models. It is generally accepted [see Cicchetti (1973)] that forecasting models based on population-specific surveys are reduced form equations which reflect characteristics of the population as well as opportunities in the form of natural resources. The basic approach of these models is to estimate and forecast activity days, defined as the number of occasions of a particular activity for a specified time for a region or for the nation as a whole.

The structure of a forecasting model can be developed from the definition of activity days, denoted v. By definition

$$v = nx, \tag{36}$$

where n = the number of users and x = trips per user. Define

$$\pi = n/P, \tag{37}$$

where π is the proportion of users in a population and P is the population size. Then (36) and (37), together, imply:

$$v = \pi P x. \tag{38}$$

Custom initiated by Davidson, Adams and Seneca (1966) and supported by empirical results has lead to the estimation of two functions, a participation rate function and a frequency of use function. The participation rate function,

$$\pi = \pi(z^1, S^1), \tag{39}$$

depends on a vector of population or household characteristics z^1 and a vector of supply characteristics S^1. The frequency of use function,

$$x = x(z^2, S^2), \tag{40}$$

may depend of different vectors of population or household characteristics and supply variables. Support for estimating separate functions is found by Deyak and Smith (1978) who show that congestion levels are more likely to influence the decision to participate in camping than the frequency of participation. Vaughan and Russell (1982a) find that variables influencing the decision to fish are different from the variables which influence the frequency of fishing. However, Heckman (1976) shows that estimating (39) and (40) separately requires maintaining the hypothesis that the random error in (39) is independent of the random error in (40).

The basic equations (39) and (40) are typically estimated from population-specific samples. Such samples are drawn from the whole population, of whom some participate in the activity in question. Thus, the participation equation (39) is estimated from a vector of binary variables where one indicates participation by the individual or household and zero indicates no participation. While binary dependent variables in the linear model create heteroscedastic errors and the potential for predictions outside the unit interval, Deyak and Smith argue convincingly that ordinary least squares as a solution is competitive with probit or logit models. The frequency of participation equation (40) is to be estimated as a conditional equation, conditioned on being a participant in the activity.

Empirical results for forecasting models are based on surveys with roughly similar designs. Surveys by the Fish and Wildlife Service, the now defunct Bureau of Outdoor Recreation of the Department of Interior, and the National Marine Fisheries Service of the Department of Commerce have periodically covered broad regions or the nation. Results from their roughly comparable findings establish some basic relationships, especially for participation models.

One of the earliest and still the most thorough piece of research on forecasting models is the work of Cicchetti (1973). Using a National Bureau of Outdoor Recreation survey, Cicchetti estimates linear versions of (39) and (40) for 24 different recreational activities, and uses the models to forecast activity days to the year 2000. Some basic relationships established by Cicchetti and corroborated by others pertain to:

(1) Age. Typically, age is an important determinant of participation, and for strenuous activities, age of participant has a curvilinear effect, first increasing and then decreasing participation, reasonably captured by a quadratic function of age.

(2) Income. Family income is frequently a statistically significant determinant of participation, in contrast to its role in demand for site functions.

(3) Supply variables. Even broad regional aggregates, such as recreational acreage, acres of water of various quality, etc. help explain participation.

The findings concerning the supply variable, while not surprising in theory, enable researchers to predict the impact of changes in recreational facilities and may eventually be used in judging policy changes.

In addition to these general relationships established by Cicchetti, other researchers have established some rather plausible relationships for participation in specific activities. Using the 1975 National Survey of Hunting, Fishing, and Wildlife Recreation, Miller and Hay (1981) show that participation in hunting in general is an inverted U-shaped function of age, and that the acres of recreational land and commercial forest in the respondent's state have a significant impact on hunting. They are unable to refute the hypothesis that income has no impact on the decision to hunt. Additional analysis of this survey [Hay and McConnell (1981)] for the decision to participate in wildlife watching and photography shows income to be important statistically, as well as a measure of the diversity of species of birds. In an analysis of the 1972 National Recreation Survey, Deyak and Smith (1978) find that various measures of camping capacity are statistically significant determinants of the decision to camp, whether developed or remote camping is considered.

A study by Vaughan and Russell (1982a) gives some insight into the use of participation models. This study, designed to estimate benefits of improving water quality, is based in part on the 1975 National Hunting, Fishing and Wildlife Recreation Survey of the Department of Interior. The research estimates benefits for three kinds of recreational fishing: cold water fishing (basically trout), warm water fishing (basically bass), and rough fish. The number of anglers fishing for the ith type of fish is given by

$$n_i = P\pi(f)\pi(i \mid f),$$

where $\pi(f)$ is the probability of any kind of fishing and $\pi(i \mid f)$ is the conditional probability of ith type of fishing, given fishing. Vaughan and Russell estimate the following logit model for the $\pi(f)$ function (p. 78):

$$\begin{aligned}\ln \pi/(1-\pi) = & -1.17 + 0.051A - 0.000075A^2 - 1.18S - 0.32M \\ & +0.46R_1 + 0.55R_2 + 0.42R_3 + 0.92I, \qquad (41)\end{aligned}$$

when

A = respondent's age,

$$S = \begin{cases} 1 & \text{female,} \\ 0 & \text{otherwise,} \end{cases}$$

$$M = \begin{cases} 1 & \text{if respondent lived in metropolitan area,} \\ 0 & \text{otherwise,} \end{cases}$$

R_i, $i = 1, 3$, dummy variables for West ($i = 1$), Central ($i = 2$) and South ($i = 3$),

I = acres of fishable freshwater/population for the respondent's state.

All of the coefficients in (41) are significantly different from zero at a very high level. And the supply variable, fishable acres per capita, is an important avenue for introducing changes in water quality. The probabilities of participating in the different types of fishing are even more closely linked to supply variables, in a way which allows one to analyze the impact of change in the different levels of water quality.

This example concludes the discussion of the forecasting models for outdoor recreation. These models are pragmatic attempts to explain behavior of outdoor recreationists with diverse backgrounds facing a wide variety of recreational opportunities. While they do not lend themselves directly to benefit estimation, they are useful tools, and as the work by Vaughan and Russell shows, these models can be used in conjunction with benefit estimates.

4. The quality of the recreational experience

In this section we focus on the attributes of recreational sites and the quality of the recreational experience. The number of trout released, the number of deer sighted, the size of crowds at a beach, the water temperature of a lake, the odor of a river, the number of hikers encountered on a trail are all aspects of the quality of the recreational experience and determine, in part, the demand for specific sites. The issue of quality is relevant for two reasons. First, efficient management of recreational facilities is partly the management of attributes of the sites, such as hiking trails, the number of campsites, etc. and requires some knowledge of how recreationists respond to these attributes. Second, some economic benefits of environmental improvements are manifested in better recreational opportunities. The larger task of measuring the otherwise unmeasured economic gains induced by improving the environment thus requires some knowledge of how the improvements change recreational behavior.

In the following analysis it is assumed that there are aspects of quality under the control of some management authority which can be measured, at least in index form. Furthermore, it is assumed that recreationists perceive and respond to these aspects. [For evidence that perceptions do not always match up to the actual quality, see Hanemann (1978).] There are obviously many aspects of quality (weather, terrain, length of day) which are not under anyone's control but this need not hinder an analysis of those which are controllable.

In modeling the quality aspect of the recreational experience, the issue of endogeneity arises. For conceptual and econometric reasons, endogeneity is important. At least three different cases of endogeneity can be differentiated. The simplest case occurs when the quality indicator is strictly exogeneous to aggregate the individual demand for the site. For example, when the recreation site is on a river and the quality is a function of sedimentation induced by upstream farming, changes in recreational behavior do not influence quality. The next case is that the quality indicator is strictly exogenous with respect to individual behavior, but is influenced by aggregate behavior. For example, the level of fish stocks in a lake can reasonably be assumed independent of the recreational activities of one individual, but is determined by aggregate recreational demand for the site. A third case, more complicated still, allows the individual himself to determine the quality indicator, through a household production function process. For example, when the number of fish caught per trip is the indicator of quality, an individual may help determine the quality by the purchase of inputs such as services of a guide, boat rentals, live bait, etc. In the third case, the quality of experience for an individual will be determined not only by his own activities but by aggregate demand as well.

Grasping the nature of the interaction between environmental quality and recreational choice is essential for two reasons. First, for econometric purposes, when quality is determined simultaneously with the choice of recreation activities, systems methods likely need to be used in estimating demand functions. Second, and more important, for purposes of policy, an endogeneous environmental quality variable cannot serve as an instrument. For example, environmental quality cannot control the levels of catch of fish for recreationists, because this level is determined in part by the anglers themselves.

The issue of endogeneity obviously becomes more complicated if one allows for the passage of time, when, for example, habits change and stocks of fish grow or decline depending on the past behavior of recreationists. In the following sections, temporal adjustment will be ignored. Furthermore, only the first two cases will be addressed: (1) congestion, the case where quality is endogenous to aggregate demand but exogenous to individual demand; and (2) environmental quality in the instances when this quality is exogenous to individual and aggregate demand.

4.1. Congestion of recreational sites

Congestion at recreational sites is an important determinant of the quality of the site because high density induces wear and tear on the facilities and because the recreational experience, especially the wilderness experience, is sought as an antidote to the crowds and congestion of everyday life. The basic model of congestion in outdoor recreation arose from a consideration of the benefits of wilderness recreation [see Fisher and Krutilla (1975)].

Congestion develops at a recreation site because of the physical relationship between the aggregate quantity demanded and the size of the facility. The measurement of congestion varies with the nature of the activity and the site. For example, congestion on beaches may be the number of bathers per acre per day. For skiing, it could be the length of lines at the lifts. In wilderness recreation, it is typically measured as the number of encounters per trip. In keeping with other sections, we assume that users are homogeneous, though admittedly part of the richness of congestion stems from heterogeneous users. Let the congestion level be denoted q.

In modeling congestion, it is essential to distinguish between individual and aggregate demand. The role of congestion on site demand can be traced from the utility function to demand for the site in the following way. Let $u(x, z, q)$ be an individual's utility function, where z is a vector of Hicksian goods, x is trips to the site, and q is the expected level of congestion at the site. For simplicity, q is expected to be the same for all days during the decision-making period. Increases in q reduce utility, x and z constant. If, in addition, q reduces the marginal utility of x and has no impact on the marginal utility of z, then $\partial x/\partial q < 0$. Aggregate demand, with homogeneous users, can be expressed as

$$v = nx = nf(c, q), \tag{42}$$

which is equivalent to (2) when there is only one zone.

Equilibrium in the aggregate requires that expectations about congestion be realized. Thus, in addition to (42), when q is expected congestion, we also need a condition such as

$$q = q(v), \tag{43}$$

where $q(v)$ is the actual level of congestion as it depends on aggregate demand.

Expressions (42) and (43) represent the equilibrium. When c is only the travel cost, and there is no entrance fee, it is open access equilibrium, and suboptimal. For each entrance price, there is a different congestion–constant demand curve, because changing the entrance price changes the level of aggregate attendance. Increases in entrance prices thus have two effects: they move each user up his congestion–constant demand curve, reducing aggregate demand, and because aggregate use is lower, congestion is lower, and the congestion–constant demand curve shifts. In Figure 15.6, when costs are c^0 and the entrance price is zero, trips are v^0 and the relevant demand schedule is $nf(c, q^0)$. When an entrance price of p is imposed, per-unit cost becomes $c^0 + p$, quantity demanded goes down, congestion falls, and the new equilibrium is moved from A^0 to A^1. The curve AB is a locus of equilibrium points but it has no normative significance.

The optimal level of use of congested recreation facilities is generally less than the open access level, because of the external costs of congestion. When a user takes one additional day, congestion goes up by one unit and all users' willingness to pay for access declines. This extra cost is ignored by each individual. Welfare improvements can be induced by charging entry prices, the optimal price being equal to the external congestion cost at equilibrium. [For a survey of the analytics of congested recreation facilities, see McConnell and Sutinen (1983).]

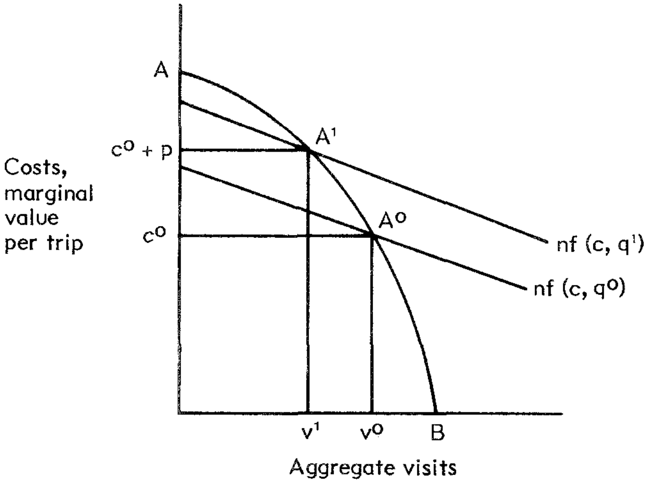

Figure 15.6. Congestion and aggregate demand.

A significant challenge is the measurement of congestion externalities. For reasons of data collection and intra-seasonal variation in congestion, this measurement is accomplished with contingent valuation methods. The work of Cicchetti and Smith (1973) is a good example. In that study, the researchers investigate the willingness of backpackers to pay for avoiding camp and trail encounters with other backpackers. Respondents gave their willingness to pay at various hypothetical levels of encounters. A statistical fit of the relationship between willingness to pay and the number of encounters and other demand function arguments shows congestion to be a significant determinant of willingness to pay. Similar results have been discovered in a variety of settings, quite different from the wilderness recreation of the Cicchetti–Smith research. The work by Walsh, Miller and Gilliam (1983) uses a contingent valuation approach to determine the costs of congestion for skiers in Colorado. Their work shows that two variables, the length of wait in the lift lines and the density of skiers on slopes, are exceptionally significant determinants of the willingness to pay for skiing.

4.2. Environmental quality and the demand for recreation

The efforts to measure the impact of environmental quality on the demand for recreation serve a much broader purpose than the efforts to measure the impact of congestion. The economic costs of foregone recreation and the economic benefits of improved recreational opportunities comprise a major part of the rationale for a policy for a cleaner environment. Thus, the costs of congestion relate to site management while the costs of environmental quality relate to the broad issue of resource management.

Both air and water pollution influence the recreational environment. Air pollution in the form of acid rain reduces the biological productivity of lakes, damaging the fishing quality and, in turn, reducing the recreational opportunities. The influence of water pollution in the form of suspended solids, biological oxygen demand (BOD), and persistent chemicals is pervasive. The recreational opportunities in countless rivers, lakes, and bays have been diminished or eliminated by water pollution. Perhaps the major role of recreation economics in the future will be to account for the economic benefits of environmental improvements. Here we review current attempts to measure the economic value of recreational opportunities created or foregone by environmental changes (see also the discussion in Chapter 6 of this Handbook).

The basic theory of the welfare change induced by changes in environmental quality is relatively simple (see Chapter 6). Let q be a measure of the environmental quality of a recreational site, for example, a physical measure such as BOD, dissolved oxygen, or simply an index of the quality. When increases in q increase the marginal utility of trips to the site, but do not affect the marginal utility of other goods, then $\partial x/\partial q > 0$. Here, we assume that q is exogenous to both individual and aggregate trips. For the individual who faces cost per trip of c, environmental quality q^0, and reservation price p^*, the value of access to the site is

$$b=\int_c^{p^*} f(p,q^0)\,\mathrm{d}p, \tag{44}$$

where $f(c,q^0)$ is the individual demand curve. When environmental quality improves (q^0 to q'), the change in site benefits given by

$$b=\int_c^{p^*} f(p,q')\,\mathrm{d}p-\int_c^{p^*} f(p,q^0)\,\mathrm{d}p \tag{45}$$

can be attributed to policies designed to improve the environment. When q is weakly complementary to x, that is, changes in q are of no value without visiting the site, then expression (45) measures the individual's willingness to pay for the change in environmental quality from q^0 to q'. By aggregating this expression over individuals and distance zones, we can compute the recreational benefits of improving the quality of the environment.

A graphical interpretation of expression (45) is useful. In Figure 15.7, the outward shift in the demand function is induced by improvements in water quality. With costs per trip fixed, the increased willingness to pay for the site is the shaded area, given by expression (45).

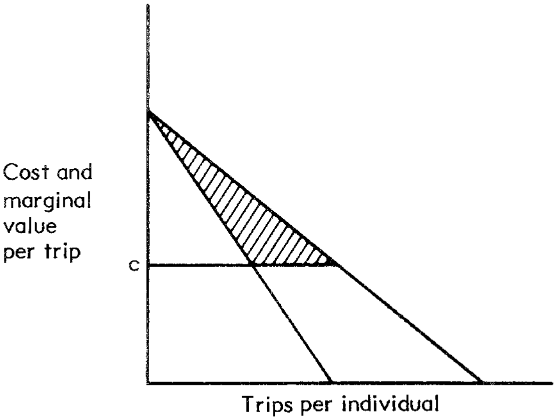

Figure 15.7. Changes in quality and individual demand.

While there are many conceptual issues dealing with valuing the increments of environmental quality, especially issues connected to endogeneity and the adjustment to quality changes, the primary issues are empirical. The basic empirical issues deal with the circumstances under which researchers are able to measure, directly or indirectly, the shaded section representing willingness to pay for environmental quality shown in Figure 15.7. As in valuing access to sites, there are two basic approaches to valuing changes in the environmental quality of sites: the travel-cost approach and the contingent valuation approach.

There is a variety of approaches for including the effect of quality variables in the travel-cost models. Approaches such as Cesario and Knetsch (1976) develop site indices. Many single-site approaches include the quality variable as an argument in the demand equation. Multiple-site models such as Burt and Brewer handle quality differences implicitly; demands for different sites need not assume that their quality is the same. Vaughan and Russell (1982b) and Desvousges, Smith and McGivney let parameters of single-site models be determined by quality variables.

Perhaps the earliest work to access the benefits of environmental improvements in recreation is a study of fishing in Yaquina Bay, Oregon by Stevens (1966). Water pollution from kraft mills influenced the level of fish stocks. Quality is measured as catch of salmon per outing, lagged one period. From time-series data on participation and catch of salmon, Stevens estimates the effect of changes in catch on the demand for trips. With cross-section data, he estimates a travel-cost model from which aggregate benefits are computed. (Here it should be noted that

Stevens uses monopoly revenue rather than consumer's surplus as a measure of benefits.) Then, assuming benefits proportional to trips, he calculates the change in benefits induced by a change in the expected catch. Although Stevens uses an inappropriate measure of benefits, his study is the first attempt to link recreational behavior with environmental quality.

The major difficulty in calculating benefits of quality improvements in a travel-cost framework is the econometric issue of estimating trips as a function of an exogenous quality variable from cross-section data sets. Most studies do not have access to time-series data as in the Stevens study. For single-site analysis within a period of time, either the quantity variable is constant or it shows short-run random or seasonal variation. In neither case will measured quality be suitable for use in a travel-cost demand function. However, while environmental quality at a single site may not vary within a time period, it certainly varies across sites. Several recent approaches have been designed to take advantage of the intersite variation in quality.

The systematic varying parameter model, initially exploited by Vaughan and Russell (1982b), is an innovative way to utilize inter-site variation. Vaughan and Russell use this approach to allow the parameters of demand for fee fishing to be determined by site characteristics.

Following Vaughan and Russell, Desvousges, Smith and McGivney (1983) adopt the systematic varying parameter model for the following site-specific travel-cost demand function:

$$x_i = \alpha_0 + \alpha_1 c_i + \alpha_2 y_i + \text{error}, \tag{46}$$

where the variables x, c, and y are trips, costs, and income. This model is estimated separately for 43 sites, yielding 43 sets of estimates for the α's. Then the α's are estimated as functions of site quality data which varies across sites. Their most convincing results are based on a 22-site subset of the original sites, when the environmental quality variables are the means and variances of dissolved oxygen. The results show that a one-unit increase in mean dissolved oxygen increases the estimate of α_0 by 0.0045 and the estimate of α_1 by 0.0002 (Table 7-14). The impact of an increase in mean dissolved oxygen is to increase the demand for trips to the site and to decrease the responsiveness of demand to cost increases, both yielding greater consumer's surplus.

The essence of the environmental quality problem from the consumer's perspective involves choosing among sites with different travel costs and different levels of environmental quality. This problem has been fully worked out by Hanemann (1978, 1982) for the case where a consumer chooses only one of a number of sites, using a discrete choice model. The discrete choice approach to modeling recreation behavior is also investigated in Feenberg and Mills (1980) and Caulkins, Bishop and Bouwes (1982).

The work on determining the impact of environmental quality on recreational benefits via shifts in the travel-cost demand function is at the frontier of recreation economics. Current work in primarily conceptual, but empirical applications include studies by Morey (1981), Brown and Mendelsohn (1983), and Samples and Bishop (1983).

The contingent valuation approach has been used many times to value the changes in the environmental quality of recreational resources. One of the earliest studies [Randall, Ives and Eastman (1974)] deals with the value of visibility to recreationists. Other studies have been similar in approach to the Randall et al. study. Respondents are presented with hypothetical variations in the environmental quality variables, and then through some kind of bidding or other response, their willingness to pay for the quality is ascertained. The contingent valuation approach is flexible enough to allow wide variations in researchers' approaches. Hence, there is no general model or approach to analyze. A more detailed study of various approaches may be found in Desvousges, Smith and McGivney (1983, ch. 4).

Environmental policy relates the value of improvements in the environment to the enhanced value of recreational opportunities. Economists have reached some consensus on the valuing of sites, but have less experience in valuing the quality of sites. Hence, future work in the economics of outdoor recreation will make a substantial investment in models which allow the measurement of the economic value from changes in environmental quality.

5. Conclusion

The economics of outdoor recreation has emerged as a different but identifiable body of literature whose object is the study of natural resources for recreational use. The distinctive characteristic of recreational activity is that the services usually are provided by the public sector. This characteristic strongly flavors the nature of research on outdoor recreation economics: it is welfare economics in action with the task of providing the public sector better information about the allocation of recreational resources.

As a tool for analyzing decisions in the public sector, outdoor recreation economics has become primarily the study of the demand for outdoor recreation. Public sector decisions about natural resources tend to be lumpy; whether to utilize a resource for a specific purpose such as recreation. Hence questions about production and pricing are only infrequently entertained. This survey has attempted to convey the nature of research in outdoor recreation economics by focusing on conceptual and empirical issues which arise in the use of the two most prevalent methods of estimating the demand for or benefits of outdoor recreation: the travel cost approach and contingent valuation approaches.

This survey has attempted to show how recreation economics has developed as a synthesis of two forces: one is the need of the public sector for information about the recreational use of natural resources; the other is the intellectual community which asks whether techniques used to determine public policy pass the test of reason. Both the travel cost method and contingent valuation approaches show evidence of this development. Travel cost models are the result of attempting to impose the axioms of choice on recalcitrant survey data sets, with econometrics providing modest success. Contingent valuation approaches owe their popularity to the ease with which they satisfy the public sector's need for benefit estimation. The continuing development of these methods is a consequence of trying to determine whether such benefit estimates accord with the axioms of choice.

A modest research agenda

As part of the conclusion of a survey of imperfect and evolving tools, it seems reasonable to suggest some areas of research which may prove fruitful for improving the usefulness of the tools. The following agenda is suggested.

(1) The value of time and substitute sites. In the travel cost method, attempts to include the value of time have linked this value to the wage rate. Two related issues worthy of interest are whether the value of time is, in fact, a parameter, and if it is not, how it is determined by the availability of substitute sites.

(2) Persistence in behavior and endogenous tastes. The ability to predict the outcome of future actions is hampered by the use of cross-section data, especially for outdoor recreation. The absence of smoothly functioning markets and the concomitant absence of information flows suggest that habits play a large role in decisions in outdoor recreation. Research into the role of habits and persistent behavior, as well as the allied concept of endogenous tastes, can make a lasting contribution to outdoor recreation economics, especially the travel cost method.

(3) Testing contingent valuation responses. Outdoor recreation economics provides a useful setting in which to test the contingent valuation approach. When this approach is used to value access to a site, it must embody information equivalent to the aggregate demand schedule for the site. A reasonable approach to testing the contingent value approach would be to use the predictive capability embodied in it to predict trips and compare actual versus predicted trips with the travel cost method.

(4) The benefits of environmental improvement. This is a broad topic, but it is clear that recreation economics will be called upon increasingly to measure the economic value of improvements in environmental assets. In this area, there is

need for advances in theory, in the sources of data, and in the estimation of models.

References

Bishop, Richard C. and Thomas A. Heberlein (1979) "Measuring Values of Extra Market Goods: Are Indirect Methods Biased?", *American Journal of Agricultural Economics* 61, 927–930.

Bishop, Richard C. and Thomas A. Heberlein (1980) *Simulated Markets, Hypothetical Markets, and Travel Cost Analysis: Alternative Methods of Estimating Outdoor Recreation Demand*, working paper 187 (University of Wisconsin, Department of Agricultural Economics).

Bishop, Richard C., Thomas A. Heberlein and Mary Jo Kealy (1983) "Contingent Valuation of Environmental Assets: Comparisons With a Simulated Market", *Natural Resources Journal* 23, 619–633.

Bockstael, Nancy E. and K.E. McConnell (1980) "Calculating Equivalent and Compensating Variation for Natural Resource Facilities", *Land Economics* 56, 56–63.

Bockstael, Nancy E. and K.E. McConnell (1981) "Theory and Estimation of the Household Production Function for Wildlife Recreation", *Journal of Environmental Economics and Management* 8, 199–214.

Bockstael, Nancy E. and K.E. McConnell (1983) "Welfare Measurement in the Household Production Framework", *American Economic Review* 83, 806–814.

Bockstael, Nancy E., Ivar Strand and Michael Hanemann (1984) *Time and Income Constraints in Recreation Demand Analysis*, working paper (Department of Agricultural and Resource Economics, University of Maryland).

Bowes, Michael D. and John B. Loomis (1980) "A Note on the Use of Travel Cost Models with Unequal Zonal Populations", *Land Economics* 56, 465–470.

Brown, Gardner M., John Charbonneau and Michael Hay (1978) *The Value of Wildlife Estimated by the Hedonic Approach*, working paper no. 6 (U.S. Fish and Wildlife Service, Washington).

Brown, Gardner M. and Robert Mendelsohn (1984) "A Hedonic-travel Cost Model", *Review of Economics and Statistics* 66, 427–433.

Brown, William G. and Farid Nawas (1973) "Impact of Aggregation on the Estimation of Outdoor Recreation Demand Functions", *American Journal of Agricultural Economics* 53, 246–249.

Brown, William G., Colin Sorhus, Bih-lian Chou-Yang and Jack A. Richards (1983) "Using Individual Observations to Estimate Recreation Demand Functions: A Caution", *American Journal of Agricultural Economics* 65, 154–157.

Brown, William G., Ajmer Singh and Emery N. Castle (1964) *An Economic Evaluation of the Oregon and Steelhead Sport Fishery*, Technical Bulletin 78 (Oregon State University Agricultural Experiment Station).

Burt, Oscar R. and Durward Brewer (1971) "Estimation of Net Social Benefits from Outdoor Recreation", *Econometrica* 39, 813–827.

Caulkins, Peter, Richard Bishop and Nicholas Bouwes (1982) "The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitute Effects", Paper presented at the American Economics Association meetings, New York.

Cesario, Frank J. (1976) "Value of Time in Recreation Benefit Studies", *Land Economics* 52, 32–41.

Cesario, Frank J. and J.L. Knetsch (1976) "A Recreation Site Demand and Benefit Estimation Model", *Regional Studies* 10, 97–104.

Cesario, Frank J. and J.L. Knetsch (1976) "Time Bias in Recreation Benefit Studies", *Water Resources Research* 6, 700–704.

Christiansen, Jens B. and Colin Price (1982) "A Note On the Use of Travel Cost Models with Unequal Zonal Populations: Comment", *Land Economics* 58, 395–399.

Cicchetti, Charles J. (1973) *Forecasting Recreation in the United States* (D.C. Health and Co., Lexington).

Cicchetti, Charles J., Anthony C. Fisher and V. Kerry Smith (1976) "An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy", *Econometrica* 44, 1259–1276.

Cicchetti, Charles J. and V. Kerry Smith (1973) "Congestion, Quality Deterioration and Optimal Use: Wilderness Recreation in the Spanish Peaks Primitive Area", *Social Science Research* 2, 15–30.

Clawson, Marion (1959) "Methods of Measuring the Demand for the Value of Outdoor Recreation", Reprint 10 (Resources for the Future, Washington).

Clawson, Marion and Jack L. Knetsch (1966) *Economics of Outdoor Recreation* (Johns Hopkins University Press for Resources for the Future, Baltimore).

Davidson, Paul, F. Gerard Adams and Joseph Seneca (1966) "The Social Value of Water Recreational Facilities Resulting from an Improvement in Water Quality: The Delaware Estuary", in: Allen V. Kneese and Stephen C. Smith (eds.), *Water Research* (Johns Hopkins University for Resources for the Future, Baltimore) pp. 175–211.

Davis, Robert K. (1963) "The Value of Outdoor Recreation: An Economic Study of the Maine Woods", Ph.D. Thesis, Harvard University, Cambridge.

Desvousges, William H., V. Kerry Smith and Mathew P. McGivney (1983) "A Comparison of Alternative Approaches for Estimating Recreation and Related Benefit of Water Quality Improvements", Research Triangle Institute for the Environmental Protection Agency, N.C.

Deyak, Timothy and V. Kerry Smith (1978) "Congestion and Participation in Outdoor Recreation: A Household Production Function Approach", *Journal of Environmental Economics and Management* 5, 63–80.

Dwyer, John F., John R. Kelley and Michael D. Bowes (1977) "Improved Procedures for Valuation of the Contribution of Recreation to National Economic Development", University of Illinois Water Resources Center, Urbana-Champaign.

Feenberg, Daniel and Edwin S. Mills (1980) *Measuring the Benefits of Water Pollution Abatement* (Academic Press, New York).

Fisher, Anthony C. and John V. Krutilla (1972) "Determination of Optimal Capacity of Resource-based Recreation Facilities", *Natural Resources Journal* 12, 417–444.

Gum, Russell L. and William E. Martin (1975) "Problems and Solutions in Estimating the Demand for and Value of Rural Outdoor Recreation", *American Journal of Agricultural Economics*, 57, 558–566.

Hammack, Judd and Gardner M. Brown, Jr. (1974) *Waterfowl and Wetlands: Toward Bioeconomic Analysis* (John Hopkins University Press for Resources for the Future, Baltimore).

Hanemann, W. Michael (1978) "A Methodological and Empirical Study of the Recreation Benefits From Water Quality Improvement", Ph.D. Dissertation, Department of Economics, Harvard University, Cambridge.

Hanemann, W. Michael (1980) "Measuring the Worth of Natural Resource Facilities: Comment", *Land Economics* 56, 482–486.

Hanemann, W. Michael (1982) "Applied Welfare Analysis with Qualitative Response Models" (Department of Agricultural and Resource Economics, University of California, Berkeley).

Haspel, Abraham E. and F. Reed Johnston (1982) "Multiple Destination Trip Bias in Recreation Benefit Estimation", *Land Economics* 58, 364–372.

Hay, Michael J. and Kenneth E. McConnell (1979) "An Analysis of Participation in Nonconsumptive Wildlife Recreation", *Land Economics* 55, 460–471.

Heckman, James (1976) "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables", *Annals of Social and Economic Measurement* 5, 475–492.

Hicks, J.R. (1956) *A Revision of Demand Theory* (Oxford University Press, London).

Knetsch, Jack L. (1963) "Outdoor Recreation Demands and Benefits", *Land Economics* 39, 387–396.

Knetsch, Jack L. (1964) "Economics of Including Recreation As a Purpose of Eastern Water Projects", *Journal of Farm Economics* 46, 1148–1157.

Knetsch, Jack L. and Robert K. Davis (1966) "Comparisons of Methods for Recreation Evaluation", in: Allen V. Kneese and Stephen C. Smith (eds.), *Water Research* (Johns Hopkins University Press for Resources for the Future, Baltimore).

McConnell, Kenneth E. (1973) "Some Problems in Estimating the Demand for Outdoor Recreation", *American Journal of Agricultural Economics* 57, 330–334.

McConnell, Kenneth E. and Nancy E. Bockstael (1984) "Aggregation in Recreation Economics: Issues of Estimation and Benefit Measurement", *Northwest Journal of Agricultural and Resource Economics* 13, 181–186.

McConnell, Kenneth E. and Ivar E. Strand (1981) "Measuring the Cost of Time in Recreation

Demand Analysis: An Application to Sportfishing", *American Journal of Agricultural Economics* 63, 153–156.

McConnell, Kenneth E. and Jon G. Sutinen (1984) "An Analysis of Congested Recreation Facilities", in: V.K. Smith (ed.), *Advances in Applied Micro-Economics* (JAI Press, Greenwich) 3, 9–36.

Miller, Jon R. and Michael J. Hay (1981) "Determinants of Hunter Participation: Duck Hunting in the Mississippi Flyway", *American Journal of Agricultural Economics* 63, 677–684.

Morey, Edward R. (1981) "The Demand for Site-specific Recreation Activities: A Characteristic Approach", *Journal of Environmental Economics and Management* 8, 345–371.

Munley, Vincent G. and V. Kerry Smith (1977) "A Note on Learning-by-doing and Willingness-to-travel", *The Annals of Regional Science* 11, 61–70.

National Park Service, U.S. Department of Interior (1949) *The Economics of Public Recreation* (GPO, Washington).

Nichols, Lee M., Marianne Bowes and John F. Dwyer (1978) *Reflecting Travel Time in Travel-cost-based Estimates of Recreation Use and Value*, Forestry Research Report no. 78-12 (Agricultural Experiment Station, University of Illinois at Urbana-Champaign).

Pollak, Robert A. and Michael Wachter (1975) "The Relevance of the Household Production and Its Implications for the Allocation of Time", *Journal of Political Economy* 83, 155–277.

Randall, Alan, Berry Ives and Clyde Eastman (1974) "Bidding Games for Valuation of Aesthetic Environmental Improvements", *Journal of Environmental Economics and Management* 1, 132–149.

Robinson, Warren D. (1967) "The Simple Economics of Public Outdoor Recreation", *Land Economics* 43, 71–83.

Samples, Karl C. and Richard C. Bishop (1983) "Estimating Social Values of Sport Caught Fish–A Suggested Approach", (Department of Agricultural Economics Working Paper, University of Wisconsin).

Seckler, David W. (1966) "On the Uses and Abuses of Economic Science in Evaluating Public Outdoor Recreation", *Land Economics* 42, 487–494.

Schulze, William D., Ralph C. d'Arge and David S. Brookshire (1981) "Valuing Environmental Commodities: Some Recent Experiments", *Land Economics* 57, 151–172.

Sellar, Christine, John R. Stoll and Jean-Paul Chavas (1982) "Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques", (Department of Agricultural Economics Working Paper, Texas A&M).

Smith, V. Kerry (1975) "Travel Cost Demand Models for Wilderness Recreation: A Problem of Non-nested Hypothesis", *Land Economics* 51, 103–110.

Smith, V. Kerry (1975) "The Estimation and Use of Models of the Demand for Outdoor Recreation", U.S. Department of Interior, *Assessing Demand for Outdoor Recreation* (U.S. Government Printing Office, Washington).

Smith, V. Kerry, William H. Desvousges and Mathew P. McGivney (1983) "The Opportunity Cost of Travel Time in Recreation Demand Models", *Land Economics* 59, 259–278.

Smith, V. Kerry and Raymond J. Kopp (1980) "The Spatial Limits of the Travel Cost Recreational Demand Model", *Land Economics* 56, 64–72.

Stevens, Joe B. (1966) "Angler Success As a Quality Determinant of Sport Fishing Recreational Values", *Transactions of the American Fisheries Society* 95, 357–362.

Strong, Elizabeth J. (1983) "A Note On the Functional Form of Travel Cost Models With Zones of Unequal Populations", *Land Economics* 59, 342–349.

U.S. Department of Interior, Fish and Wildlife Service (1982) *1980 National Survey of Hunting, Fishing, and Wildlife-Associated Recreation* (U.S. GPO, Washington).

Vaughan, William J. and Clifford S. Russell (1982a) *Freshwater Recreational Fishing—The National Benefits of Water Pollution Control* (Resources for the Future, Washington).

Vaughan, William J. and Clifford S. Russell (1982b) "Valuing a Fishing Day: An Application of a Systematic Varying Parameter Model", *Land Economics* 58, 450–463.

Vaughan, William J., Clifford S. Russell and Michael Hazilla (1982) "A Note On the Use of Travel Cost Models With Unequal Zonal Populations: Comment", *Land Economics* 58, 400–407.

Walsh, Richard G., Nicole P. Miller and Lynde O. Gilliam (1983) "Congestion and Willingness to Pay for Expansion of Skiing Capacity", *Land Economics* 59, 195–210.

Willig, Robert D. (1976) "Consumer's Surplus Without Apology", *American Economic Review* 66, 589–597.

Ziemer, Rob F., Wesley N. Musser and R. Carter Hill (1980) "Recreation Demand Equations: Functional Form and Consumer Surplus", *American Journal of Agricultural Economics* 62, 136–141.

Chapter 16

ECONOMICS OF ENVIRONMENT AND RENEWABLE RESOURCES IN SOCIALIST SYSTEMS

Part 1: Russia

MARSHALL I. GOLDMAN

Wellesley College

PART 2: CHINA

SHIGETO TSURU

Hitotsubashi University, Tokyo

PART 1: RUSSIA

1. Introduction: Environmental myth and reality in a socialist system

It is widely agreed, and repeatedly emphasized in this Handbook, that industrial pollution is a major source of negative economic externalities. Air, water, and solid residuals from private production can easily be directed downwind, downstream, and into landfills in such a way that the private producer bears only a small share and on occasion none of the disposal costs. Instead, those costs are absorbed by others downwind or downstream. These costs thus can be seen as a consequence of a system in which private producers and consumers manage to avoid costs which economic efficiency demands should be theirs to bear. Instead, the polluters push off these costs onto society as a whole.[1] The solution to this situation, some have argued, would be to nationalize all industry.[2] As such critics see it, once private industry is eliminated, there will no longer be an incentive to generate external costs. In contrast, by definition, in a system of state-owned industry where all the means of production are owned by the state, there is no difference between social and private costs. Thus, society as a whole becomes the

---

[1] A.C. Pigou, *Wealth and Welfare* (MacMillan and Co., London, 1912) p. 159; Oscar Lange and Fred M. Taylor, *The Economic Theory and Socialism* (University of Minnesota Press, Minneapolis; 1938) pp. 103–104; Evegeni K. Fedorov, *Man and Nature* (International Publishers, New York, 1982) p. 76.

[2] Barry Commoner, *The Closing Circle* (Alfred A. Knopf, New York, 1971) pp. 274–279.

Handbook of Natural Resource and Energy Economics, vol. II, edited by A.V. Kneese and J.L. Sweeney

beneficiary and loser in all economic activity. Consequently, in socialism, it might be argued, there will be no temptation to pollute since it is society as a whole that enjoys the gains of production but also the burdens of disposing of residuals.

Since all the means of production are owned by the State in the Soviet Union, presumably there are no private costs, no private–social costs dichotomy, and so effective control of pollution.[3] Thus, for most economic theorists outside the Soviet Union it comes as quite a surprise to discover just how serious the pollution problem is in the U.S.S.R. This is all the more surprising given the fact that the Soviet Union has enacted some of the most far-reaching pollution control laws in the world.[4]

These laws anticipate almost all possibilities and spell out very strict environmental controls. Yet the very fact that the Soviet authorities have enacted such laws suggests that state ownership by itself is no guarantee that state-production enterprises will operate any differently than private-production enterprises. If state-production enterprises did take it upon themselves to absorb all social costs, presumably there would be no need to introduce regulatory legislation.

In the pages that follow, we shall examine what there is in the Soviet system which either serves to facilitate or prevent pollution. First we shall see to what extent theory and practice diverge and to what extent the country's culture and endowment have led to or away from pollution. Second, we shall discuss how the lack of pluralism and the overwhelming emphasis on production make it difficult to protect the Soviet environment. Third, we shall examine how, if at all, Marxist theory facilitates pollution. Fourth, we shall consider to what extent the Soviet incentive system, whether or not it is a Marxist one, can be a cause of pollution. Fifth, we shall consider the impact of the Soviet emphasis on secrecy on the environment. The next section of the chapter will briefly examine resource policy in the Soviet Union, and finally, we will see where the Soviet system has advantages in dealing with pollution.

2. Why theory is not practice: Culture and endowment

The Soviet experience indicates that Soviet environmental theory and law are more impressive on paper than in practice. The Soviet Union has long had legislation setting limits on waste discharge that is by far more demanding than any other country in the world.[5] Yet, it was widely acknowledged in the U.S.S.R. that these limits were set more to impress the outside world than to constrain Soviet producers. Such laws also serve to beguile Soviet authorities: "We have the

---

[3] I. Petrianov-Sokolov, *Literaturnaia gazeta*, 24 November 1971, p. 13.

[4] Philip R. Pryde, "The 'Decade of the Environment' in the U.S.S.R.", *Science* 220, no. 4594 (15 April 1983) p. 279, footnotes 22 and 23.

[5] Marshall I. Goldman, *The Spoils of Progress: Environmental Pollution in the Soviet Union* (MIT Press, Cambridge, 1972) pp. 25–26.

laws, therefore we have no problems." Moreover, pollution is widely claimed to be a capitalist problem, and for a long time, the official position was that there is little or no pollution in the Soviet Union.[6] On occasion, when it was impossible to deny the reality of some major incident, it would be acknowledged as an aberration. More often, state authorities would take steps to ban all discussion or acknowledgment of pollution incidents. If no one knew, then by implication the problem did not exist. Thus, in the early 1970's, after Western critics first began to pinpoint pollution problems in the Soviet Union, Soviet authorities imposed censorship on the discussion of pollution incidents in the Soviet Union.[7] Unfortunately, the prohibition on discussion of pollution proved to be more effective than the ban on pollution itself.

Much of the reason for the existence of pollution in the Soviet Union arises from the fact that the Soviet Union has placed and continues to place so much emphasis on economic growth. In that sense, it tends to resemble a developing country, and it may be that the behavior patterns of a developing country dominate whatever socialist influences that the Soviet Union might have. Whenever the drive for economic growth tends to prevail over all else, whether the country be rich or poor, environmental control will not come easily. Efforts to control pollution usually necessitate additional investment or operating costs. For a developing country, this is usually a significant burden relative to the country's total wealth.

More often than not, funds set aside for pollution control are funds that could otherwise be used for more traditional forms of economic growth. In a sense, pollution control is a luxury, and thus a state which has committed itself to rapid economic growth tends to concentrate its resources on growth and to postpone or avoid diversions from growth such as would be associated with pollution control.

This helps to explain the tendency in the Soviet Union, contrary to socialist theory, to impose social costs on individual citizens rather than on state enterprises. The effects of this tendency extend well beyond pollution phenomena. In the field of consumer good distribution, for example, state planning officials traditionally have restricted the flow of investment funds which in other societies would be allocated for the construction of shops and retail store outlets. This is regarded as a low priority expenditure in the Soviet Union. Soviet planners have concluded that the funds can be better spent elsewhere. As a result, Soviet consumers have become inured to standing in long lines for the vast majority of products they buy. By holding down the number of retail outlets, Soviet planners have forced the Soviet consumer to spend more time in line. The human capital expended for this purpose tends to be much larger than is the case in nonsocialist states.

---

[6] *Current Digest of the Soviet Press*, 17 July 1968, p. 16; Boris Kamarov, *The Destruction of Nature in the Soviet Union* (Sharpe Publishers, White Plains, N.Y., 1980) p. 17.

[7] Kamarov, p. 16

But, when analyzing pollution control in the Soviet Union, it is also necessary to allow for the historical and cultural heritage of the country. For a pure theorist, this may seem to be a strange thing to do, but it is impossible to evaluate behavior in the Soviet Union without references to the past. When something is done in the Soviet Union, it is very difficult to judge whether the motivation is a consequence of socialist ideology, expediency in the pursuit of economic growth, or Russian tradition and heritage. For example, the vastness of the land mass, and its huge store of natural resources, encompassed within the borders of the Soviet Union have had a major impact on how the Soviet authorities treat their natural endowment. With so much land at their disposal, Soviet planners and ministry managers tend to take a causal attitude toward the exploitation of their land and resources. What difference does it make if the water is polluted, the arable land flooded, or the resources exploited? Given so much land and space, it is hard to challenge such attitudes. Moreover, history seems to support the idea that Russia's resources are limitless. Thus, periodically there have been fears that Russia would run out of oil. As early as 1900, when such doubts were first expressed, drillers managed to find new fields. When the oil fields were depleted in the Baku region, new fields were discovered in the Northern Caucusus.[8] These fields in turn were supplanted by fields in the Volga–Ural region which in turn were supplemented by the giant fields in West Siberia.[9]

Against the backdrop of such natural wealth, it is not altogether a surprise that Soviet officials do not normally concern themselves with the husbanding of the country's natural resources. Thus, the relative level of wholesale prices of the Soviet natural resources tend to be significantly lower than comparable prices in the noncommunist world. Soviet economists frequently have complained that raw material prices are too low, especially for fuels.[10]

The economists were concerned in part because these relatively low prices put a premium on heavy utilization of the country's resources. Soviet manufacturers have become accustomed to paying prices for raw materials that are often below the cost of production. Indeed, energy prices have been understated to such an extent that Soviet coal enterprises operated at a loss from 1977 to 1981, and even the Soviet oil industry operated at a declining rate of profit.[11] While the petroleum industry's rate of profitability as a percentage of capital stock was as high as 27.8 percent in 1970, it fell below 8.6 percent in 1981.

---

[8] Marshall I. Goldman, *The Enigma of Soviet Petroleum: Half Empty or Half Full?* (Allen & Unwin, Winchester, Ma., 1980) p. 19.

[9] Goldman, *Enigma*, p. 119.

[10] *Ekonomicheskaia gazeta* no. 17, April 1980, p. 7; V.E. Orlov, "O nekotorykh voprosakh ekonomiki otraslei toplivno-syr' evogo i energeticheskogo kompleksa", *Finansy SSSR*, August 1979, pp. 19–20; A. Komin, "Zadacha sovershevstvovaniia optovykh tsen i tarifov v promyshlennosti", *Planovoe khoziaistvo*, May 1980, p. 34.

[11] Tsentral'noe statisticheskoe upravlenie S.S.S.R., *Narodnoe khoziaistvo S.S.S.R.*, 1922–1982 (Finansy statistika, 1982, Moscow) p. 552 [hereafter referred to as Narkhoz with the appropriate year]; Narkhoz, p. 506.

The whole incentive system with its emphasis on meeting production targets and low prices for natural resources has generated one of the most resource-intensive economic systems that exist. Not surprisingly, Soviet products tend to be very heavy. Thus, Soviet engines use more metal per unit of horsepower and more fuel per ton of steel than is true of most capitalist countries. As a corollary, the Soviet Union has had very limited success in trying to stimulate miniaturization.

It is interesting to speculate about the kind of economic pricing and incentive system that might have evolved if the first communist country had not been established in the Soviet Union but in a resource-scarce region like Japan or Switzerland. There does not seem to be anything in the nineteenth-century Marxist literature to suggest that the Soviet type of pricing and incentive system is the only one permissible. Indeed, Marx left few guidelines of any sort as to how to operate a socialist system. Presumably, if the Soviet state had been located in a country with few resources, more attention would have been paid to the value of the resources, and for that matter, to the environment. Under these circumstances, the result would then have been relatively higher prices for raw materials than those that have actually evolved in the Soviet Union. The point is, that in examining whether socialism is good or bad for environmental control, we have to take into consideration not only what economic theory might tell us, but also the resource endowment and history of the country being studied.

3. Production above all when there is no pluralism

Regardless of the culture and the endowment of a Socialist state, there are other factors in the nature of socialism itself that tend to increase the risk of environmental disruption. One that is often overlooked arises from the political consequence that there is usually little if any pluralism in a communist state like the Soviet Union. In the Soviet Union, this has meant that the state has vigorously opposed the creation of any other institutionalized power base not directly under the control of the government. Thus, there is no outside power base to complain about the way the state is being run or the way industry is being operated, since industry itself is an organ of the government. As a result, state organs both govern and are governed. Occasionally, *ad hoc* groups may come together now and then temporarily to stand watch on a specific issue, but there can be no independent or persistent watchdog.

This is not to deny that the state seeks to foster the formation of groups concerned with the environment. The best example is the Society for Protection of Nature. This is a group unlike anything else in the world. It has over nineteen million members. But the Soviet Union's trade union organization is also one of the largest in the world (except perhaps for China). Like Soviet trade unions, however, the Society for the Protection of Nature does as it is bid. Its president for many years was Nikolai Ovsiannikov, who also happened to be the first

Deputy Minister of the Ministry of Land Reclamation and Water Management of the Russian-Soviet Federation of Socialist Republics (RSFSR), a Ministry with one of the worst environmental records in the Soviet Union. Clearly, such environmental groups are designed to implement state goals, not challenge them. If anything, they seek to co-opt environmental protests and diffuse them.

It is interesting to compare the situation in the democratic world. Certainly, the situation there works imperfectly, but the existence of pluralistic forces frequently works to curb pollution. In many instances, the government finds itself cast in the role of referee or judge. On one side of the government stands the manufacturer or other potential polluter, and on the other side the environmentalist. Depending on the relative strength of the political pressure currently being generated, the state may move to curb industrial emissions or to curb environmental protests. The Soviet government seldom finds itself cast in such an intermediary role. The state, as we have seen, owns all the industry and therefore is the industrial polluter. As long as it is determined that increased production is the main priority, one state organ is unlikely to challenge another, particularly when the polluter is the manufacturer.

Gosplan is the chosen instrument for achieving economic growth, and little if anything is ever allowed to challenge that emphasis. Indeed, anything that impedes growth is suspect. Thus, stress on growth affects not only the behavior of Gosplan, but of individual factory managers and even the workers. Their bonuses, premiums, and promotions are dependent on whether or not the factory fulfills its quantitative production targets. They do not take lightly any diversions from the production process. Consequently, there is enormous pressure to insure that all the investment funds made available to the factory are put to work increasing production, not to diversionary expenditure. Not surprisingly, there are numerous complaints that funds originally intended for pollution control are diverted instead to increased production.[12] This is another illustration of the tendency even in a socialist system to externalize the social costs onto society as a whole.

All of this, however, does not mean that it is impossible for environmental risk to become a public issue. One of the most dramatic instances of the battle between those who want Soviet industry to bear its social costs and those who want the environment at large to absorb these costs, was fought on the shores of Lake Baikal. The fight attracted international attention – in part because there seemed to be so much concern in the Soviet Union itself from unauthorized critics. This in itself was something of note. Another reason for the concern was the uniqueness of the lake itself, which in turn explains why there seemed to be such determination to challenge the economic planners and ministry officials. As I have pointed out, this is an uncommon phenomenon.

---

[12] *Pravda*, 3 August 1968, p. 1; *Kazakhstanskaia pravda*, 17 June 1973, p. 3.

What made the lake so unusual? Many factors! It is probably the oldest lake on the planet. It is just over a mile deep, which makes it nearly three times deeper than Crater lake, the deepest lake in the United States. It is the largest body of fresh water in the world. The mineral content of the water is 50 to 25 percent lower than most other nonsalt lakes. In other words, the water quality approaches that of distilled water. This is due in part to the nature of the rock in the watershed. The quality of the rock is such that it results in the release of little in the way of dissolved minerals or suspended matter into the inflow to the lake. Similarly, the cold temperatures hold down the evaporation which in turn holds down the buildup of salts. In this unique environment, it is not surprising to discover that the lake contains over seven hundred living organisms that are unique to it, including the world's only fresh water seals and the *golomyanka* – a transparent fish whose young are born alive.

The uniqueness of the lake, particularly the low mineral content of the water, made the lake a tempting target for the Soviet paper and cellulose industry. A paper, and especially a cellulose, plant benefits from large quantities of high-grade water. To the extent that Lake Baikal water could be utilized, there would be no need to process it beforehand. That would reduce the direct cost to the factory. In addition, to the extent that the plant could also discharge and dilute its wastes into the vast depths of the lake, it would be spared the direct costs of building a water treatment plant. The fact that the effluent was of poor quality did not initially deter those who proposed the plant's construction. To the planner, the availability of a cheap pure source of water supply as well as the ability to discharge waste into that water was what was important. They were not overly concerned that this would impose on society, or at least nature, the costs that might otherwise have to be carried by the Ministry of Pulp and Paper.

The pressure to avoid what should be individual plant operating costs by shunting them all onto Soviet society as a whole is not limited to Gosplan and Ministry executives. These sentiments are occasionally shared by scientists as well. Thus, P.L. Kapitsa, an Academician and one of the Soviet Union's most respected scientists and a person not known for his anti-environmental stance, has offered the following rationalization for using Lake Baikal as a recipient of effluent:

> The value of Lake Baikal does not lie simply in the abundance of clean water, but in the fact that it functions as a biological filter of tremendous capacity producing clean water. The water that enters the lake is much dirtier than the water that emerges. This purification is the result of biological processes in Baikal.... The industrial significance of Lake Baikal lies in the fact that it represents a huge purifier of water, and our concern should be directed toward preserving that capability. It is therefore wrong to say, "Don't touch Lake Baikal." The lake should be exploited, but so as not to disturb its life processes

> or interfere with its water-purifying properties. We must therefore know how and to what extent Baikal may be polluted so that it may continue to process dirty water and yield clean water.[13]

The plans of the Ministry of Pulp and Paper were discussed and publicized by limnologists and naturalists working with the lake. At first, their warnings in 1960 were directed to residents in the immediate vicinity of the lake, but by 1962, articles began to appear in the national press. The uproar that followed was virtually unprecedented. Some of the Soviet Union's leading writers and scholars sent off a barrage of letters, for the most part in uncoordinated fashion, complaining about the avariciousness of the Ministry of Pulp and Paper Industry.

The campaign was partially successful. Ultimately, a high-level advisory commission was created to judge charges of malfeasance, and a combined session of Gosplan, the Presidium of the Academy of Sciences, and the Committee on Science and Technology was convened to decide upon remedies.[14]

That this effort accomplished anything was most unusual given the dominance of the drive toward economic growth and industrial expansion. The Ministry of Pulp and Paper did not agree to cancel construction plans, however, nor did it agree to build what would have been an unusually expensive sewage pipe that would have extended forty-two miles over mountainous terrain, at the cost of at least forty-two million dollars. This was deemed by the experts to be the only treatment method that would have insured removal of paper plant wastes to regions outside the Lake Baikal watershed. Instead, the Ministry of Pulp and Paper argued that it would be enough to build an advanced water treatment facility, despite the fact that the residual from this treatment plant was then discharged into the lake.

Encroachment on the lake is now no longer limited to the emissions of the Ministry Pulp and Paper Industry plant. Recently, reports have been published indicating that air pollution from the factories in the Baikal Basin is also causing problems.[15] Similarly, G.I. Glazii has warned that a whole series of industrial enterprises are evading the intent of governmental laws. Instead, they are discharging waste into Lake Baikal; again, to spare themselves social treatment costs.[16] Potentially more threatening is the plan to build the Kholodnaia Lead and Zinc Refinery on the northern shore of the lake. This could have the most destructive impact of all. Yet again, the prospect of the earnings from that ore is very tempting, especially if the refinery can be built in such a way that some of the emissions can be released without full treatment.

---

[13] Pryde, p. 275.

[14] Thane Gustafson, "The New Soviet Environmental Program; Do the Soviets Really Mean Business?", *Public Policy* 26, no. 3 (Summer 1978) p. 474.

[15] *Ekonomicheskaia gazeta* no. 49 (December 1978) p. 13; *Current Digest of the Soviet Press*, 10 September 1980, p. 9.

[16] *Current Digest of the Soviet Press*, 7 July 1982, p. 1.

That even the defenders of the unique and fragile Lake Baikal have such a hard time illustrates just how powerful and relentless Gosplan and the industrial ministries have become. What is impressive is that the environment receives even as much attention as it does. It is not easy to override Gosplan. First of all, Gosplan sits at the peak of the Soviet power structure. The Director of Gosplan is also Deputy Chairman of the Council of Ministers, and equally important, Gosplan determines the resource allocation patterns for the various Ministries. Moreover, Gosplan and many of the industrial ministries are situated at the All Union level. This is significant because most of the environmental laws of the Soviet Union have been passed at the Republic level. This makes it hard to call All Union entities to account.

The hierarchical pattern of Gosplan and the All Union Industrial Ministries is not an accident. Their structural position reflects not only the priority attached to economic growth in the Soviet Union but the fact that leadership is uninterested in having another agency with equal powers that might chose to check or challenge any of Gosplan's actions. For a time, the Ministry of Health was entrusted with the promulgation and enforcement of environmental concerns. The Ministry of Health was a member of the Council of Ministers and theoretically could at least question the actions of other ministries and Gosplan itself. But the Ministry of Health tended to restrict itself to matters of health. Moreover, it may have been a duly constituted member of the Council of Ministers, but it lacked the clout of the industrial ministries. Thus, it was relatively ineffective when it came to environmental concerns which did not have a direct impact on health. For that matter, there were instances where pollution did involve health concerns, and even then the Ministry of Health sometimes found that when it began to insist on enforcement, its officials ran the risk of being fired.[17]

In retrospect, the choice of the Ministry of Health as the pollution watchdog was not necessarily illogical. After all, for many years the United States Public Health Service was assigned similar responsibilities for some aspects of pollution control and prevention. But as the pollution problem mounted in magnitude and complexity, it became clear that, at least in the United States, a more powerful and more comprehensive organization was needed. This became urgent particularly when the nonhealth aspects of pollution became more and more of a problem, and the whole nature of the issue became more complex. Eventually, the United States Congress decided to create a new multifaceted entity that was not subordinated to one or a narrow set of jurisdictions. The result was the Environmental Protection Agency (EPA) and the Council on Environmental Quality (CEQ), both of which were endowed with significant powers and reported directly to the President. This meant that they were no longer subject to a conflict of interest when dealing with the Department of Interior or the Department of Health, Education and Welfare. There were virtually no governmental agencies in

---

[17] *Meditsinskaia gazeta*, 14 April 1983, p. 4.

the Executive Branch that could ignore the EPA and CEQ. Moreover, the requirement that almost all government agencies would have to file Environmental Impact Statements with the CEQ before they could embark on any major program had the effect of placing significant influence in the hands of the CEQ.

The early success of the EPA and CEQ led several countries around the world to create similar agencies with similar powers. The Soviet Union was not one of them. The main reason for the lack of action by the Soviet Union was that if an EPA-like organization with real power had been set up, by definition it would have been able to impose constraints on Soviet production activities. Gosplan, and for that matter Soviet leaders, were not prepared for such interference or diversions from the economic growth objective.

The matter came to a head in 1972. In May of that year, the United States and the U.S.S.R. signed a joint pollution cooperation agreement. Ultimately, it turned out that of all the joint agreements that were signed during what was the most enthusiastic stage of the Detente Era, the environmental agreement was the most successful. The projects were more forthcoming and less political than virtually any other such exchange. But at the time, Soviet authorities had a dilemma. They had no counterpart to the EPA and CEQ. The Ministry of Health was clearly inappropriate. Did they dare create a comparable agency with comparable powers?

Several Soviet economists and environmentalists concerned about the powerlessness and ineffectiveness of the existing government watchdog agencies have called for just such a new organization.[18] The economist P.G. Oldak, in particular, has urged the formation of a new central body that would report directly to the Supreme Soviet of the Soviet Union (what passes for the equivalent of Parliament). In particular, such an organization would seek to coordinate environmental efforts. In support of this argument, Academician A.G. Aganbegian reported that try as it might, the Ministry of Agriculture has found that because there is no all powerful authority dealing with environmental matters, it could do nothing to stop the contamination of farm land from air pollution by a nearby steel mill.[19]

The creation of a powerful environmental agency might suit the American predilections, but it would sooner or later cause problems in the Soviet Union in the form of challenges and restraints on production. Reflecting such concerns, the Soviet leadership instead decided to upgrade an existing organization, the Hydrometeorological Service. Heretofore, it had been mainly a data collection agency with little in the way of enforcement power; nor has it been given much power subsequently. But it was involved with both water and air concerns, and therefore it seems superficially at least to match up with EPA. Even though the formal title

---

[18] *Current Digest of the Soviet Press*, 7 July 1982, p. 5.

[19] *Current Digest of the Soviet Press*, 7 July 1982, p. 5.

of the organization has been changed – the original name Hydrometeorological Service was upgraded in 1978 to U.S.S.R. State Committee for Hydrometeorology and Control of the Environment – but the function and political importance are about the same as before.[20]

4. The distortions of Marxist economic theory

The environmental deterioration resulting from the unwillingness or inability to temper the drive for economic growth is compounded by the fact that many of the economic mechanisms in the Soviet Union function in such a way that they produce incentives inimical to both economic efficiency and environmental protection. For example, because of ideological reasons, Soviet authorities have difficulty authorizing any reasonable land use charge. In effect, land is treated as a free good. For that reason, land use is more a function of political power than economic rationality. Those with the most influence or power tend to prevail, regardless of what the opportunity cost might be. One of the most common illustrations in the U.S.S.R. is the way agricultural land is appropriated and then flooded for use as a reservoir behind a hydroelectric station.[21]

Soviet planners from the time of Lenin have placed a high premium on building up the country's electrical capacity, preferably a capacity generated by hydroelectric power. Given the political priority and the fact that land as such is treated basically as a free good in Marxist ideology, there is no mechanism in the Soviet Union to signal that something of value is being sacrificed. Inevitably, large quantities of farm land have been and continue to be lost forever for crop growing without any real accounting for or consideration of the opportunity costs.

A water project of enormous scale is being contemplated to bring additional water supplies to Soviet Central Asia. The area is arid to begin with. Making matters worse, almost all of the few rivers in the area have been diverted upstream for use as a source of irrigation. That means that those downstream often find supplies previously available severely depleted. This was largely the reason for the falling levels of the Caspian Sea and the imminent evaporation of the Aral Sea. Compounding matters, the Central Asians argue that if they only had more water, they could substantially increase the output of cotton, a most valuable export crop that thrives in the climate of the region.

The prospect of a large increase in cotton production has captured the imagination of the planners in Moscow. However, such visions hold absolutely no fascination for residents and defenders of Siberia. They acknowledge that their region is well endowed with water, but that is not reason enough to reverse the

---

[20] *Pravda*, 31 March 1978, p. 1.

[21] *Current Digest of the Soviet Press*, 7 July 1982, p. 2.

flow of the large number of rivers that flow north from Siberia to the Arctic Ocean.[22] The Siberians argue that in order to reverse the flow of rivers toward the South, the country will have to spend at least 20 billion rubles and that in the process, valuable land will have to be flooded in the North. Moreover, despite numerous studies and debates, no one can guarantee that there will be no ecological impact on Siberia. Nonetheless, the effort to redirect the course of the rivers seems to be moving ahead.[23]

Admittedly, there is no guarantee that projects of this sort would be ruled out if prices were more meaningful and land values were more properly provided for. The projected increase in cotton production and even grain production would be worth several billions of rubles a year. Moreover, increased agricultural output would probably mean increased exports yielding hard currency, and in any event, lower hard currency agricultural imports. Thus, there would be considerable interest in pursuing such projects even if the state could draw up a meaningful cost–benefit analysis. Yet the temptation to embark on projects of this nature is all the greater given the undervaluation of the land.

Part of the Soviet government's attraction to such projects stems also from the engineering challenge involved. Soviet leaders have frequently been absorbed by the romance of "remaking nature". Such attitudes are not necessarily unique to the Soviet Union or to communism. Americans have had their own fantasies, although they apparently have now abandoned their own attempt to reroute Canadian rivers south to the United States.

The fascination with engineering and the attempt to prove how effective Soviet managers can be helps to explain some of the damage the Soviet Union has inflicted on the environment. Thus, this penchant for grandiose approaches largely accounts for the onslaught Soviet fishing authorities have made on ocean fishing grounds. The concept of mother and satellite fishing fleets is a good example of where engineering has moved ahead of nature. These fleets allow Soviet sailors to remain at sea for long periods of time in waters distant from the servicing facilities that are needed by most fishing fleets elsewhere. This means that Soviet fishermen can fish for longer periods of time in a more thorough fashion. Unfortunately, it almost appears at times that the Soviets have vacuumed the area once they have moved through. This not only leaves little for other fishermen who might follow, but because the fish catch is so large, the long-run breeding prospects of the fishing grounds are sometimes jeopardized.[24] Nor do the Soviets limit such practices only to international water bodies. They have also denuded inland water bodies inside the Soviet Union and all but cleaned them out.[25]

---

[22] *Radio Liberty*, 6 April 1982, p. 1.

[23] *Trud*, 7 April 1983, p. 4; *Liternaturnaia gazeta*, 10 March 1983, p. 11.

[24] *New York Times*, 24 July 1975, p. 1; *Pravda*, 15 December 1971, p. 3.

[25] *Pravda*, 15 April 1978, p. 2; *Radio Liberty* 228, no. 78 (18 October 1978) p. 1; *Izvestiia*, 7 August 1974, p. 3; *Pravda*, 12 January 1972, p. 2; *Literaturnaia gazeta*, 17 June 1970, p. 11.

5. Even when it works well it may work badly

While economic mechanisms in the Soviet Union work poorly when they work at all, that does not mean they have no impact on the environment. Unfortunately, it often turns out that it is a perverse impact. In part, these problems are the result of Marxist theory. As mentioned earlier, the underpricing of land leads to the possibly unwise destruction of what everyone agrees is valuable farm land. Given the ten billion dollars or so in hard currency that the Soviet Union spends every year on food imports, good agricultural land is not a free good.

The distortions are further compounded by the way Soviet prices are formulated. It is not enough that Marxist theory burdens the system with certain shortcomings, but in addition, the price-setting process works inefficiently and poorly at best. For example, the wholesale prices of most raw materials were held constant from 1967 to 1982. The net result was that the Soviet price system did not reflect changing scarcity relations for a fifteen-year period. Moreover, this was a time of far-reaching change. The failure to reflect these changes only added to the way the price system served to stimulate and indeed subsidize the needless consumption of many raw materials with accompanying residuals generation.

In part, the reason why prices are changed so infrequently is that it is a very difficult task to adjust a whole pricing system by fiat. There are simply too many interrelationships, and those relationships change too frequently. Thus, it is all but impossible to expect that all the necessary relations can be measured and provided for and that the relative weighting system once adopted will last for long. Moreover, the longer the interval between changes, the more items there must be that have to be adjusted and the harder such changes are to make. It becomes an enormous chore for Soviet planners to work out all the various relationships. Soviet economists recognize this. When asked in 1983 whether or not fifteen years was not too long a period of time in between price adjustments, a Soviet planner replied that, "Yes, it was. Henceforth we plan to adjust prices every *five* years instead!" (emphasis added).

The lack of meaningful prices and other economic stimuli reinforce the tendency among Soviet planners to resort to central decrees to bring about change. Occasionally, noncommunist government officials will attempt similar campaigns which rely on exhortation more than economic stimulus. The effort in most of the Western world to promote energy conservation above and beyond the stimulus provided by higher prices was a prime example. But in the Soviet Union, as in most communist countries, change seldom comes from the periphery. Instead, it must come from the center by means of a campaign. Khrushchev ordered that virtually the whole Soviet Union switch from wheat to corn, that a whole new territory, the Virgin Lands, be transformed into agricultural production, and that a new incentive or organizational structure be adopted by Soviet industry. Inevitably, such orders lead to all kinds of abuses. For example, to avoid antagonizing the Moscow authorities, the peasants in the northern part of Siberia in such remote cities as Uhkta decided to grow corn. With not much more than two months growing season, the result was a disaster.

Not surprisingly, such Draconian decisions will sometimes have an adverse environmental impact. For example, in yet another effort to increase agricultural output, Khrushchev concluded that the country needed to increase its fertilizer output. To do that, it naturally needed a chemical industry. Such overnight efforts, particularly when dealing with chemicals, were bound to have some unanticipated results. There was no awareness of the need for nor time to prepare for environmental safeguards.

Another campaign, this time a project by Brezhnev, involved increasing agricultural cultivation in the northern and central nonblack earth regions of the country. To do this, it was necessary to drain a significant number of swampy areas. Unfortunately, this has caused increased flooding on existing arable lands and given rise to dust storms in some of the newly drained areas. Under the best of circumstances, it is hard to insure that planners at the center will be able to anticipate all the contingencies. When campaigns of this sort come in such rapid-fire order, it is very difficult to anticipate or to avoid some untoward consequences.

6. State security and the implications for the environment

When explaining the usual reasons for the existence of environmental problems in the Soviet Union, Western analysts have focused primarily on economic and political factors. Some of the studies written by Soviet environmental specialists, especially when they write for the samizdat, or unofficial publications, stress other factors as well. One of those factors that Western observers seemed to have missed is that environmental protection is made particularly difficult in the Soviet Union because of the Soviet Union's enormous concern for secrecy and security.[26]

All societies have their security codes, but in an authoritarian, one-party state like the Soviet Union, the impact on the environment is much greater than we are normally accustomed to. Certainly, Western scholars of the Soviet Union are aware of the importance of security and secrecy in the Soviet Union, but their normal concern is with the political and economic implications. Until recently, there has been little realization that secrecy might be a significant cause of environmental degradation.

Rigorous control over news makes it very difficult to generate the public outrage that is necessary to produce the public pressure that usually underlies environmental activism and protest. What you do not know, you cannot respond to. Moreover, public protest and activism in most of the noncommunist world

---

[26] Kamarov, p. 33.

arises not so much in response to a single incident but to a series of events, which makes good news coverage all the more important.

One of the few instances where the national Soviet public was informed was in the case of Lake Baikal. This, as we saw, illustrates that public protest of this sort can be useful. To the state authorities, however, it also suggests how inconvenient the dissemination of such news can be. Not surprisingly, therefore, the Soviets try to limit access to news about environmental difficulties to officially authorized groups. This explains why the best analysis of the dimension of the environmental problem has had to come from the West.[27]

Recently, someone who at one time was a Soviet citizen wrote a samizidat or underground study of the environmental problem in the Soviet Union. While Boris Kamarov, the nom de plume of the author, does bring to light several incidents that heretofore have been missed in the West, it turns out that because of the limited information he had in the Soviet Union, he has also been victimized by several mistakes that were actually easier to catch outside the Soviet Union. In addition, the relatively small size of his study reflects the fact that the Soviet policy of secrecy limited the size of his data base.

But what makes Kamarov's study most interesting is the access he had to some pollution incidents that were unknown outside a very limited circle within the Soviet Union. Some of the secrecy was due to the Soviet authorities' concern for military security. It is in areas related to military activities where the insistence on security is particularly keen. Thus, according to Kamarov, for many years Soviet officials in the Ministry of Health denied that anyone in the Soviet Union was dumping polychlorinated biphenyls (PCBs) in the Baltic Sea.[28] For that matter, replied the Minister of Health, no one in the Soviet Union had ever produced PCBs. Such denials did not dissuade Swedish scientists from complaining that they had discovered large concentrations of PCBs near the mouth of the Neman River. According to Kamarov, officials in the Soviet Ministry of Health were not lying. They had checked, but as far as they knew, or anyone would let them know, the Soviet Union did indeed lack the capability of producing PCBs. What happened, in fact, was that Soviet industry did know how to produce PCBs. Soviet industry had been producing PCBs since at least the 1950s, but for most of that time, production was allocated entirely to military needs. Soviet authorities did not want anyone to learn that they had such capabilities, and therefore anyone who inquired, including the Ministry of Health, was told no such product was being produced. Ultimately, Kamarov claims that Soviet environmental officials began to conduct laboratory tests for themselves and discovered PCBs in water bodies scattered throughout the Soviet Union, including Lake Baikal.

The reluctance to acknowledge the existence of PCBs and the preoccupation with secrecy suggests that the Soviet Union will likely have more difficulty than

---

[27] Philip Pryde, *Conservation in the Soviet Union* (Cambridge University Press, 1972).

[28] Kamarov, p. 32.

other industrial countries in dealing with the phenomenon of hazardous wastes.[29] So far, no one country in the world seems to have handled hazardous waste disposal with success. This is in part because waste dumping was the accepted disposal method for many years, and many dump sites were never systematically registered. Thus, after a time there would be no memory of what was dumped and who dumped it. Added to this, it often happened that private dumpers did their dumping in secret to avoid having to pay for the cleanup costs. But for the most part, environmentalists in the West seldom have to worry that state authorities like those in the U.S.S.R. would do all they could to insure even additional secrecy.

It is striking that so far there seems to have been relatively little if any discussion about hazardous waste disposal in the Soviet Union. As the PCB case illustrates, however, the lack of public discussion about hazardous waste is not to be taken as a sign that the problem does not exist in the U.S.S.R. Only rarely is the problem even alluded to. In the newspaper *Meditsinskaia gazeta* (*Medical Journal*), 14 April 1982, page 4, its correspondent, Vaitseva, complains about a "sludge pond" in the city of Tula. Since the sludge consists of residuals from a ferrous metal combine, the likelihood is that this is a hazardous waste problem, even if the concept of "hazardous waste" is still all but unknown in the Soviet Union.

In part, such limited attention to this hazardous waste disposal problem as such is due to the fact that it has been only recently that American authorities have come to realize the extent to which hazardous waste disposal can be dangerous. When a major environmental problem is uncovered in the United States, it often takes Soviet officials some time before they discover that it also exists in the Soviet Union. Yet, the odds are high that the Soviet Union has comparable difficulties. If anything, the hazardous waste problem is compounded by the fact that the State is reluctant not only to acknowledge that it has such problems, but also that it may be producing something with military potential.

There is very good reason to believe that the treatment of nuclear energy and nuclear waste disposal in the Soviet Union reflects the same syndrome. Soviet authorities have committed themselves to building up their nuclear energy output. Indeed, during the Eleventh Five-Year Plan, almost all new electrical generating capacity in the western part of the country was to come from the construction of nuclear generating facilities. By 1985, the Soviets hope to have nuclear power stations produce 14 percent of all the Soviet Union's electricity.[30] What is equally remarkable about this decision is that it has provoked almost no popular protest. Soviet authorities have somehow managed to convince the Soviet people that the

---

[29] Kamarov, p. 32.

[30] Lesley J. Fox, "Soviet Policy in the Development of Nuclear Power in Eastern Europe", in: Joint Economic Committee, *Soviet Economy in the 1980's: Problems and Prospects*, Part 1 (U.S. GPO, Washington, D.C., 1982) p. 487.

problems associated with nuclear energy are about the same as, if not less than, those of fossil fuel energy production. Soviet scientists repeatedly argue that "environmental pollution by harmful emissions (including radioactive emissions) is many times greater from thermopower stations than it is from atomic power stations. Thus, the widespread development of atomic power engineering is justified not only from an economic standpoint, but also from an ecological one."[31]

Equally impressive, the Soviet scientific community seems very passive about the whole affair. In part, this is a result of a carefully orchestrated campaign led by nuclear scientists in the Soviet Union who tend to be firm believers in the safety of their program. In fact, A. Aleksandrov, the chairman of the Academy of Sciences and therefore one of the country's most prestigious officials, has been particularly outspoken about the advantages and safety of nuclear energy. He brushes aside the anti-nuclear protests in the West as a diversion promoted by the monopolists who control fossil fuel extraction and power generation, especially the multinational oil concerns.[32] "The development of large nuclear power stations could endanger the profitability of fuel-producing monopolies." That is why, according to Aleksandrov, American newspapers "exaggerated out of all proportion" the incident at Three Mile Island.

In contrast, because there are no private monopolies in the Soviet Union, Soviet officials assert that they have no comparable problem. After all, Soviet scientists would not lend themselves to any such exercise. The very fact that Soviet scientists build such facilities show that nuclear energy is safe. If it were not safe, Soviet scientists would never build such facilities, and since they build nuclear power stations, they obviously must be safe. This is said without much of any qualification, despite the fact that there have been convincing reports of a serious nuclear accident near Kyshtyma in the Urals in 1957–58.[33] Apparently, this incident was a consequence of poor nuclear processing technology. That seems to be why when a rare doubt is raised about nuclear energy in the Soviet Union, it tends to be focused almost entirely on a concern about the adequacy of Soviet nuclear waste treatment.[34] Of course, this involves not only nuclear energy waste disposal, but nuclear weapons waste disposal as well.

But there is every reason to believe that there may be more problems ahead. Reportedly, one of the nuclear reactors the Soviets have sold to Finland developed a crack. Since the Soviets tend to put more care into export items than those intended for the domestic market, presumably the Finnish reactor was one of their better efforts. If there were quality problems with the foreign model, it is safe to assume that there have probably been at least comparable problems at

---

[31] *Soviet News*, 15 May 1979, p. 150; *Pravda*, 4 June 1980, p. 2.

[32] *Soviet News*, 15 May 1979, p. 150.

[33] *The New York Times*, 23 April 1979, p. 15.

[34] *Pravda*, 7 August 1979, p. 3.

home. If that is true, that could prove to be very troublesome, since until the Three Mile Island incident, the Soviets were not in the habit of building containment vessels over their reactors.[35]

7. Resource policy

In the same vein, the way the Soviets treat renewable and nonrenewable resources leaves much to be desired. In large part, these difficulties are a natural consequence of the Soviet price system. As we have also seen, the price system is hampered by the fact that it bears the defects imposed on it by a residual obeisance to Marxist theory. Thus, land, and by extension timber, as well as minerals in the ground are undervalued to the extent that they are accorded any value at all.[36]

Admittedly, the fact that the State, not private owners, controls all of the Soviet Union's mineral rights means that the Soviet Union has been spared one type of overpumping of petroleum that frequently plagues oil drilling in the unregulated capitalist countries. Unlike the early Texas oil drillers, for example, there is no need in the Soviet Union for one driller to attempt to extract petroleum from the oil pool which also extends under his neighbor's property line. Pirating of this sort tends to lead each of the capitalist neighbors to steal from the other's oil field which in turn leads to a needlessly overzealous pumping of the entire pool. Such overpumping usually causes a premature drop in gas pressure in the deposit which in turn results in a lower overall extraction of oil than if it had been pumped gradually. In contrast, in the Soviet Union, the whole pool, not just the portion directly under the driller, belongs to the State. Therefore, there is no need to try to outpump your neighbor. Yet, it turns out that the end result is the same, private property or no private property. The similarity is due to the fact that Soviet managers find themselves not in a race to outproduce their neighbors, but to exceed the production targets assigned to them by the State. To qualify for a premium, the manager must fulfill, or even better, overfulfill, the yearly or at best the five-year plan targets. In effect, these are short-run goals when what is called for are goals set to maximize the maximum efficient rate of recovery (MER) of each petroleum basin. The MER may require a five-year time horizon, but it is more likely to require a longer period of time. Consequently, the setting of annual or five-year goals means that overall output per basin will not achieve the maximum output possible.

Ironically, authorities assigned to regulate oil production in most of the states of the United States have developed new control methods which serve to bring about increased basin production. Admittedly, this is an infringement on private

---

[35] *The New York Times*, 15 January 1980, p. A6.

[36] *Current Digest of the Soviet Press*, 7 July 1982, p. 6.

property rights. In contrast, Soviet oil producers continue to race the plan so that Soviet economists find themselves complaining that Soviet oil fields actually yield a lower rate of petroleum extraction than comparable American fields.[37] This is a complaint that holds for almost all mineral extraction activities in the Soviet Union.[38]

Soviet minerals are not only extracted inefficiently, but the exploration process itself is generally far less efficient than it might be. One of the most blatant examples of this inefficiency is the incentive system used by crews which explore for petroleum fields. Like almost everyone else in the Soviet production, Soviet drilling crews are judged in quantitative terms.[39] In this case, this means the number of meters drilled can be measured more precisely than the virtually unmeasurable "increase in reserves". However, emphasis on number of meters drilled as often as not does not lead to any increase in reserves. The reason is that drillers generally discover that the best way to maximize the number of meters drilled is to drill only short holes. The deeper the hole, the more likely it is that the drill pipe will break, which means the more time it will take to replace and reinsert a longer chain of pipe, which means the longer the drilling process will take, which means the lower the work bonus will be. Since they don't have to worry about finding any actual reserves, it is therefore not surprising that Soviet drill teams prefer to drill short holes in already proven areas where the service facilities and housing amenities are already established and therefore are more comfortable than they would be in as yet unexplored areas where the new and unknown fields might be.[40] What seems particularly hard to understand, however, is that all of this goes on despite the fact that Soviet authorities are well aware of the distortions that result and have been criticizing this procedure for well over five years with little to indicate that any corrective measures have been taken.[41]

We have discussed at some length how the Soviet price system does not reflect the scarcity relationships of Soviet raw materials. As long as this is the case, not only will mineral extraction be ineffective, but also the raw material consumption. To some extent, these distortions have been partially corrected by the wholesale price reforms of January 1982. However, the impression continues to exist that in the eyes of the Soviet consumer, the price of Soviet raw materials tends to be seriously understated. In the case of energy, for example, the low price more than anything else explains why the Soviet Union has one of the world's highest ratios of change of energy usage to change in GNP, the so-called energy elasticity coefficient.

---

[37] T. Khachaturov, "Prirodnye resursy i planirovania narodnogo khoziaistva", *Voprosy ekonomiki* (August 1973) p. 17.

[38] N. Feitl'man, "Ekonomicheskaia otsenka prirodnykh resursov", *Voprosy ekonomiki*, no. 10 (October 1980) p. 72; P. Poletaiev, "Perspektivy razvitiia okhrany priroda", *Planovoe khoziaistvo*, no. 1 (January 1982) p. 46.

[39] *Pravda*, 1 February 1983, p. 2.

[40] *Pravda*, 1 February 1983, p. 2.

[41] *Turkmenskaia iskra*, 7, December 1977, p. 2.

Because the Soviets measure GNP differently than we do in the West, it is difficult to draw up a precise energy elasticity coefficient. However, Soviet statistical authorities publish a figure each year which purports to show the change in all forms of energy consumption over the past year.[42] This figure is not universally accepted outside the Soviet Union, but it seems as good as those drawn up by Western observers. If this change in the use of Soviet energy is compared to the Central Intelligence Agency's (CIA) estimate of change in Soviet GNP as defined by Western measurements, the ratio for the period 1980–81 was a relatively high 0.89. For earlier years, it was almost always more than one. This improvement not only reflects administrative action by the state, but a response to the belated decision to double the retail price of gasoline in September 1981. This retail price increase together with the wholesale price reform of January 1982 was an important change. These adjustments have helped, but even today it appears that Soviet drivers and energy users are being subsidized to use energy.

In sum, the Soviet planning system seems no better equipped to solve what seem to be the unsolvable problems in the market and pluralistic systems of the capitalist world. For example, when one interest conflicts with another, the Soviets have no guaranteed way to satisfy both disparate parties. Similarly, despite their emphasis on increasing supplies, the Soviets have found that they are facing more and more supply constraints. They can dig deeper and further out, but more and more they find themselves exhausting their raw materials deposits and rediscovering the law of diminishing returns. Nature can only be pushed so far, and after a time, it makes no difference if the pusher is a capitalist or a communist.

8. When the system works well

Having concentrated so much on Soviet shortcomings in dealing with pollution and natural resources, it might appear as if the Soviets could do nothing right. If that is the resulting impression, it is important to point out that that was not what was intended. Soviet authorities do care about the environment – after all, they must also breathe the air and drink the water. The problem is that normally they care even more about economic growth. Thus, when Soviet officials do take decisive action, they usually do so because they have concluded that the net social cost of failing to act will exceed the net direct savings which would accrue from avoiding such expenditures. Thus, the government decided that something had to be done to save the Volga because failure to have done so would have cost even more. On 17 March 1972, the Soviet Council of Ministers adopted a resolution

---

[42] Narkhoz, 1982, p. 70.

which in effect ordered that approximately one billion rubles should be spent to finance industrial and municipal waste water treatment facilities in order to clean up the river. The fact that the Soviet Union decided to spend so much money is not surprising given that the Volga is a vital source of fresh water as well as the main sewer line for almost the whole of Central and Eastern U.S.S.R. What is surprising is that the results have been so mixed so far.[43] In other words, there are strong pressures that work to divert the resources set aside for pollution control and to use those that manage to find their way for environmental treatment at less than the full potential.

Almost everywhere the results are less than fully satisfactory. Certainly the air around Moscow and the Moskva river are much cleaner than they used to be. But Moscow continues to have serious smog problems despite the tendency to use more and more natural gas, a cleaner source of fuel, in boilers and for other heating purposes. Similarly, the Moskva is much cleaner than it used to be. The fishermen have returned, and more important as an index, so have the fish. Undoubtedly the interest in environmental protection is greater, and in many instances actual environmental conditions are better, in 1983 than they were in 1970. Yet, even so, I saw raw effluent being discharged into the Moskva River as late as January 1983. The protest over Lake Baikal and the sincere concern of many Soviet officials did help to prevent the very worst from happening. Similarly, various enterprises and ministries are now required to set aside a fixed percentage of their investment allocation for pollution control, as much as five percent, according to some reports.[44] Certainly that is a hopeful sign. However, the problems persist. So far, it appears that effort to control pollution is no more, and in many instances, less successful in the Soviet Union than it is in the capitalist world.

9. Conclusion

Conceivably, the theorists could be correct; it may be that a communist state may someday bridge the gap between private and social costs and thereby impose greater control on pollution. In addition, it may just be that the Soviet Union may not be the best model of a socialist system. Yet, because reality, at least Soviet reality, has deviated so far from the promise, the Soviet model must continue to be regarded more with skepticism than hope.

---

[43] *Pravda*, 10 August 1981, p. 3.

[44] Central Intelligence Agency, "Sluggish Soviet Steel Industry Holds Down Economic Growth", in: Joint Economic Committee, *Soviet Economy in the 1980's: Problems and Prospects*, Part 1 (U.S. GPO, Washington, D.C., 1982) p. 210.

PART 2: CHINA

1. The People's Republic of China (hereafter to be referred to as "China"), is still a developing country, with her per capita gross national product estimated to be around one-fifth of that of the U.S.S.R. This must mean that the tenet of rapid economic growth occupies a position of high priority in the conduct of Chinese national policies, inherently even much more so than in the U.S.S.R, and one is tempted to surmise that environmental concern may still be a "luxury" there.

However, China does offer the picture of a country which attempted, at least in one sphere, to pursue simultaneously a policy of economic development and environmental protection. This was epitomized in the principle of the comprehensive use of wastes, known as "the recovery and re-use of *the three wastes*, i.e. waste liquid, waste gas and waste slag". "The Three Wastes" movement was started as a nationwide campaign from the latter part of the 1960s, and it did achieve certain results in recycling of waste materials both in industry and households.[1]

Mr. Tang Ke, China's head delegate to the Stockholm Conference on Human Environment in 1972, enunciated Chinese philosophy on environment in the following manner:

> There is doubtless a possibility of industrial development causing untoward effects on environment. However, this type of problem is soluble as society makes progress and science and modern technology respond to the challenge. Just as we need not hesitate to eat simply on the ground that we might suffocate in the process, we need not slow down the tempo of our industrialization on the ground of possible accompanying environmental damages.
>
> Our government has adopted a set of policies with emphases on (1) overall planning, (2) rational allocation, (3) comprehensive utilization, (4) turning the harmful into the beneficial, (5) reliance on people's initiative, (6) cooperation by all concerned, (7) protection of environment, and (8) improving of people's welfare.[2]

Following these words, Mr. Tang referred to the then ongoing "The Three Wastes" movement.

In 1973, Chinese government held the first nationwide conference on protection of environment, and it was there that Premier Zhou Enlai proposed what is known as "The Three Principles of Simultaneity". The principles entailed (1)

---

[1] See K. William Kapp, "'Recycling' in Contemporary China", *Kyklos* XXVII, Fasc. 2 (1974) pp. 286–303.

[2] Each one of these eight items was spelled out in four Chinese characters, totaling thirty-two. Thus, the guideline was called "The thirty-two Character Principles".

designing of antipollution measures simultaneously, (2) constructing of antipollution equipments simultaneously with the construction of industrial plants, and (3) operating of antipollution equipments simultaneously with the operation of industrial plants. It was made clear at the time that unless these principles were strictly abided by, the planning board would not ratify the construction plan nor would furnish the plant site, and the financing authorities would not provide funds for the construction purpose. In actual practice, however, "The Three Principles of Simultaneity" are said to have been sidestepped more often than not during the period of dominant influence in "The Gang of Four" (until around 1976).[3]

2. The orthodox Chinese position on the causes of environmental disruption has been, as can be expected, basically Marxist. At the Stockholm Conference, a Chinese delegate pointed to "monopoly capitalist groups" as the real culprits who "discharge at will and in disregard of the fate of the people harmful substances that pollute and poison the environment".[4] In other words, the capitalist mode of production is to blame – the position which has been expatiated repeatedly by various experts in China.[5]

It follows from this basic position that under socialism, as in China, benefits and costs are not evaluated in terms of market values (said to be the outcome of "anarchic production relations of capitalism") but rather in socio-economic terms or social use values. Emphasis is placed at the same time on the relevance of the welfare of future generations in the social benefit–cost calculations related to environment and resources.

How far such basic orientation bore practical results in terms of waste recycling and pollution control is probably yet too early to judge. But observations offered both by foreign visitors and the Chinese themselves[6] seem to agree, in general, that real improvements in the environmental field were slow to come until after the promulgation of the Environment Protection Law (EPL) (Trial Enforcement) in September 1979. In other words, putting the EPL into effect was the turning point.

3. In the Constitution of the People's Republic of China, adopted in March 1978, Article XI, Paragraph 3 states that "the state protects environment and natural

---

[3] Yoshihiro Nomura, "The Environment Law", in: Ichiro Kato (ed.), *Modernization and Legal Reforms in China* [in Japanese] (Tokyo University Press, 1980) p. 131.

[4] See *Peking Review* 27 June 1972.

[5] Kapp, pp. 297–298.

[6] Cf. Zhang Wei, "Environment Protection in Beijing", *Jimin Chugoku* [in Japanese] (November 1982) pp. 79–81. In this article, the author cites quite candidly a number of concrete examples of environment offenders in and around Beijing which had continued or actually even worsened during the decade of the 1970s.

resources, and it also prevents and eliminates pollutions and other public nuisances".[7] On the basis of this constitutional prescription, the Environment Protection Law was drafted and the Fifth National People's Congress adopted it in September 1979 with a proviso that its enforcement be regarded as a trial with possible future revisions. The law sets forth the basic framework for environmental protection, subsequently to be buttressed by specific laws and regulations applicable to several aspects of environment such as air pollution, water pollution, noise pollution, etc.

Problem areas covered by the term "environment" in this basic law (Article III) are quite broad, encompassing ambient air, water, land, mineral resources, forests, grassland, wild flora and fauna, aquatic life, places of scenic and historic interest, tourist zones, hotspring and health resorts, nature conservation areas, human habitats, etc. And the guideline spelled out in the law (Article IV) was exactly the same "Thirty-two Character Principles" which were enunciated by Mr. Tang Ke in Stockholm in 1972. "The Three Principles of Simultaneity", originally proposed by Premier Zhou Enlai, was also incorporated in the law (Article VI) as a part of responsibility for individual firms.

It is noteworthy that in this connection, micro-responsibility for firms[8] included the presenting of environment assessment reports and also the adoption of the "Polluters Pay Principle" with a combination of "carrots" for the accomplishment and "sticks" for failure. These requirements were subsequently further elaborated in greater detail, specifying the kinds of "major construction projects" for which the assessment report is mandatory and the items on which *ex ante* assessments are required. On the "charges" collection also, the standard threshold levels of emission of pollutants were specified in 1983[9] relating them to graduated "charges" to be paid by offending firms. Delay in attaining the standard threshold, as well as delinquency in payment, were to be penalized through imposition of pro rata higher fines.

These innovations are in line with the apparent determination of the Chinese authorities to make the EPL (Trial Enforcement) of 1979 work effectively, as spelled out in "The Decision Relating to the Strengthening of Environment Protection in the Readjustment Period of the National Economy" adopted in February 1981. This "Decision", it is said, was formulated on the basis of a year-and-a-half experience after the passage of the EPL (Trial Enforcement) and laid special stress on the need for the improvement of technological levels of environmental management within the sphere of activities of individual firms,

---

[7] The English expression "public nuisances" is used in this translation for "kogai," which originated in Japan, and was apparently used for the first time in the official Chinese legal document.

[8] One of the unique new features of the EPL is a provision (Article XXXII) calling *an individual* to account in the cases of especially serious environmental damages.

[9] Put into effect from 1 July 1983.

highlighting more clearly than earlier the standpoint of interlinking the problem of energy and resources-saving with that of environmental protection.

The test for all these innovations is yet to come. And although one hesitates to pass judgment on China's success or failure in environmental control as a socialist state, one is tempted to conclude, at this interim state, that China, although a developing country, is more seriously concerned than the U.S.S.R. with environmental problems as a whole.

INDEX