HCP/M8429-01

MASTER

AN ECONOMIC ANALYSIS OF THE CONSERVATION OF DEPLETABLE NATURAL RESOURCES

Final Report

August 1978

Prepared for
U.S. DEPARTMENT OF ENERGY
Assistant Secretary for Conservation
and Solar Applications
Division of Data Analysis

Under Contract No. EM-75-C-01-8429

Available from:

National Technical Information Service (NTIS)
U.S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22161

Price: Printed copy: $12.00
Microfiche: $3.00

HCP/M8429-01

UC-95 a,c

AN ECONOMIC ANALYSIS OF THE CONSERVATION OF DEPLETABLE NATURAL RESOURCES

Prepared by
Micro Economic Associates
Berkeley, California

6553
950

Final Report

August 1978

For
U.S. DEPARTMENT OF ENERGY
Assistant Secretary for Conservation
and Solar Applications
Division of Data Analysis
Washington, DC 20545

Under Contract No. EM-75-C-01-8429

TABLE OF CONTENTS

III. TECHNICAL APPENDIX cont. Page

I. Summary and Conclusions

1. Introduction

This is the final report of Microeconomic Associates under contract number CO-04-50325-00 to study the intertemporal allocation of natural resources arising in a market economy, and in particular to develop parameterizations by which bias in the intertemporal allocation may be estimated.

The first interim report surveyed existing work in this area and noted the directions in which research seemed particularly needed. Work on two of these areas, the effects of imperfect competition and of the tax system, was begun in the first interim report and continued through the second. In addition, work on the consequences of imperfect risk markets was begun in the second interim report and continued in the present report. A second new area investigated in the second report was that of technical change, an assessment of market inducements for technical change and its consequences under different institutional arrangements. The second report also began the third phase of the study, the analysis of alternative policies, with an examination in detail of the BTU tax.

This report includes analysis of the effects of imperfect risk markets and two studies that deal with issues of trade policy. One study analyzes the optimal trade policy of a large consuming nation facing competitive resource suppliers. The other examines trade policies under the threat of embargoes by resource cartels, one of the important motivations for Project Independence. A concept of flexibility is introduced and we analyze the optimum degree of flexibility when the supply of a natural resource is affected by the probability of a large price rise from an embargo. Finally, the nature of policies which should be introduced in the face of the threat of an embargo are discussed.

We have been pleased with the progress we have made under this contract. Major developments in at least seven areas should be noted:

(1) The work done in conjunction with this contract represents the first comprehensive analysis of the behavior of a natural resource cartel. The results reported in the first two interim reports were surprising and unexpected.

(2) The work described in the second interim report represents the first analytical comparison of the dynamic incentives for research under monopoly and competition with the socially optimal allocation. Although some results confirmed the conventional wisdom in this area, others did not.

(3) The first and third reports contain a comprehensive analysis of taxation of natural resources, with again several results which counter conventional wisdom.

(4) The work presented in the second report constitutes the first analytical treatment of the role of monopoly in destabilizing prices.

(5) The second and third reports contain the first qualitative (and analytical) analysis of a stock market with more than one commodity; the parameters which determine the market bias in the intertemporal allocation of oil are also identified.

(6) The work discussed in the third report represents the first analytical study of the implications of "flexibility" for the intertemporal allocation of a natural resource.

(7) The work contained in the third report is the first analytical treatment of optimal tariff policy for consumption of a natural resource. The results are markedly different from conventional optimal tariff analysis.

In addition, we succeed in developing in the context of all the problems examined parameterizations which should enable the quantification of the magnitudes of the biases introduced by the various distortions analyzed.

Many of these analytical results have immediate and important policy implications and we have attempted, both in the appendices and in part I of this report, to draw out these implications. In other areas, the analysis has formed the basis for further theoretical and empirical research. Throughout our reports, we have attempted to identify areas in which further theoretical work is required, and areas where final judgments rest on empirical data. In addition to the seven major findings discussed above, this study program has succeeded in developing a consistent theoretical structure that can be used for the analysis of important policy issues within the areas of natural resources and energy. The first interim report outlined the economic theory of exhaustible resource allocation over time under perfectly competitive conditions. The analysis presented in the reports has taken the competitive economy as a benchmark, identified important areas of market failure and estimated their consequences. This work is primarily neither an exercise in computational techniques nor a theoretical treatise. Rather we have in this program developed and extended basic economic methods and applied them to important policy issues. Even in areas where at the present state of knowledge unambiguous policy recommendations cannot be made, it is clear that the considerations raised have potentially highly important ramifications.

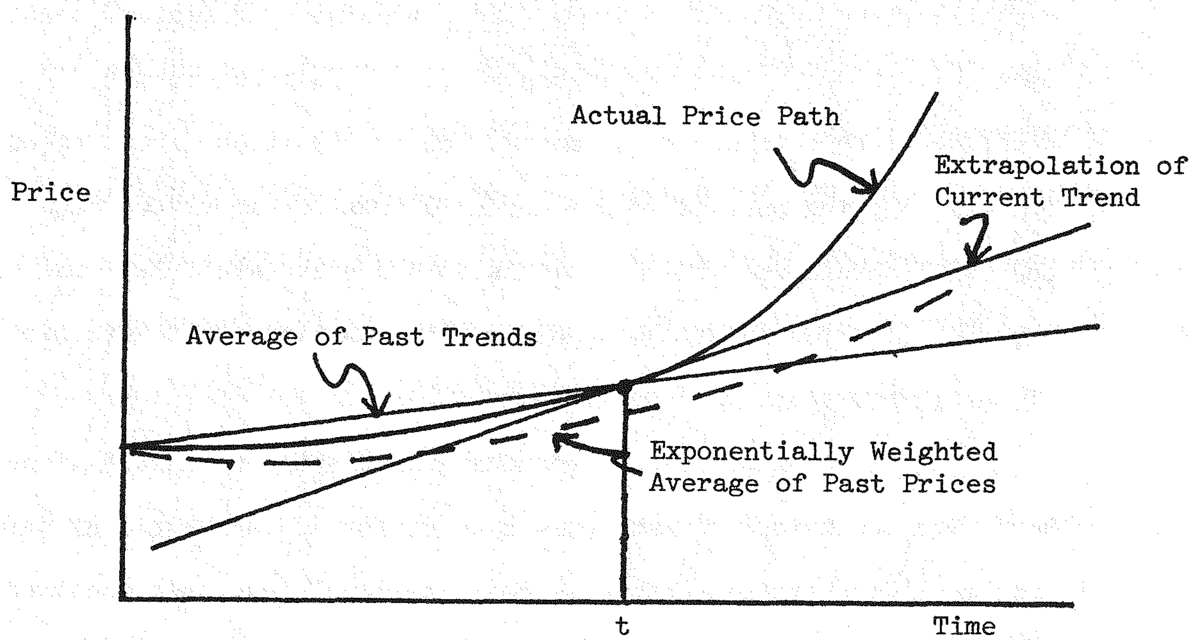
In our judgement, it would be extremely worthwhile to integrate the methodology developed in the course of this program in the quantitative work of the Federal Energy Administration. While this would require some investment of time and research effort, we feel that the end result would be a more consistent and robust framework for specific policy recommendations.

2. Investment in Energy Conservation

It is clear that a rising relative price for energy encourages the substitution of other factors for energy -- i.e. encourages energy conservation. We inquire here whether there might be reasons for supposing this market response to be inadequate: it seems that in fact there are two reasons why this might be so, each associated with a different type of market failure. On the one hand the future price of energy is not known with certainty because future markets fail to exist, and in their absence investors must make forecasts which we shall argue will tend to underestimate the future price of energy. On the other hand firm's investment decisions do not take place in a perfectly competitive undistorted market environment, but are influenced by the tax system, and the existence of price controls. Both kinds of market failure can be mitigated by government policies; it is natural to address the first problem by providing the basis for better forecasts through indicative planning, and to correct the second through a range of specially designed fiscal instruments. In part III we examine the effects of accelerated depreciation allowances, tax credits, tax exempt bond finance and loan guarantees in offsetting the incentive to underinvest in energy conservation. Indicative planning is discussed at the end of this section.

The effects of both kinds of market failure are more severe for durable equipment, and it is clear that many investment decisions in energy-using industries are, at least for a period, irreversible. Typically there is a range of technologies from which a choice may be made, and which differ in their energy-intensities, so that ex ante there is a possibility of factor substitution. However, any given technology is characterized by relatively fixed factor proportions, so that once the investment decision is implemented there is no such scope for substitution. The user is locked into a particular pattern of factor intensities for the life of the plant, which may often be several decades. In such a situation, it is particularly important that the choice of factor intensities be made, not in the light of prevailing factor prices, but in the light of those expected to rule over the life of the plant -- and it is also important that these expectations should be as accurate as possible. An illustration of this point is that many of our current problems are aggravated by the fact that much of the present capital stock is appropriate to the cheap-energy decades of the 1950's and 1960's.

In general, of course, businessmen recognize that the choice of factor intensity must be appropriate to the relative prices expected to prevail over the entire life of the plant: a problem arises because there is a likelihood that these expectations will be systematically biassed. The point is that if energy is provided from an exhaustible resource, its relative price will increase over time. Investors at time t_1 will need to predict the course of energy prices over the period t_1 to t_2 in order to choose the technology appropriate to a plant of life $(t_2 - t_1)$, and the principal evidence available to assist them in doing this is the course of prices up to t_1. A reasonable assumption is that they use some simple rule for extrapolating this -- either simple linear extrapolation of the current trend, or of an average of past trends, or perhaps exponential extrapolation. It is easy to show any of these, if there are fixed or rising costs of energy extraction, which leads to a systematic <u>underestimate</u> of future energy prices -- figure 1 illustrates.



<u>Figure 1</u>

The diagram establishes that there may be a systematic underestimation of the future increases in energy prices, so that investors will systematically underestimate the incentive to conserve energy and substitute alternative factors. In other words, private entrepreneurs are likely in practice to commit themselves to more liberal patterns of energy use than they would if they had accurate knowledge of the future.

This inefficiency which arises from a lack of accurate information about future prices is a general one, and applies to energy supply decisions as well as energy conservation investments, and to all other decisions that have to be made long in advance of the time when their consequences will become apparent. The classical examples of such decisions are those involved in planting forests or laying down services, though in modern economies they arise more commonly in the guise of decisions about investment projects with long gestations periods. In order to make such a decision well -- that is, to achieve to the greatest extent possible the objectives being pursued -- one needs information about the conditions that will obtain when the decision's consequences are felt: one would expect a positive relationship between the quality of the decision and that of the information.

In an idealized market economy, there would be institutions that would provide this information -- namely, forward markets. By establishing prices at which future contracts could be concluded, they would provide all of the information needed for national long-term decision-making. Unfortunately it is the case that for a variety of reasons modern industrial economies are poorly endowed with forward markets, with those that do exist extending only a relatively short distance into the future. This obvious shortcoming in the institutional framework of the modern market economy has led to several developments whose function is at least in part compensatory: long-term contracts are of course one such development (though their partial nature makes them of limited value), and indicative planning is another. In the following sections we analyze the nature and purposes of indicative planning with particular reference to the energy market.

Indicative Planning for Energy

Indicative planning is not easily defined. In what is certainly the major theoretical work in the field, Meade [1973] views it simply as a process which is a substitute for forward markets: in a sufficiently small community this process would involve the government convening meetings of all of those concerned, and then acting as an auctioneer and determining equilibrium prices for future trading. In practice such an approach is not feasible, and indicative planning would take the form of a planning agency forecasting the likely future development of the industrial sector concerned, and then discussing this forecast with those whose behavior will affect it or be affected by it. In the light of the discussions the forecast is revised, a new round of discussions ensues, and so on. The eventual outcome should be a forecast consistent with the actions of all and embodying the best information available: it is this forecast that is referred to as an indicative plan. It should be emphasized that it is not usually envisaged that such a plan will be implemented by fiat: companies would be free to take whatever notice they choose of it in making their decisions, but it would be hoped that it would be so evidently superior to any forecast of their own that they would to a very substantial extent be guided by it. Of course, if all decision-makers accept the plan as an accurate picture of the future and act as they said they would during the discussions, the plan will in fact have an important self-fulfilling property.

An important consideration which deserves explicit mention is the treatment of uncertainty. Any economic agent is clearly uncertain about many aspects of the future, and it is helpful to break the uncertainty, facing a decision-maker in the no-plan situation into two categories:

(i) exogenous uncertainty -- uncertainty about events whose outcome is not influenced by the operation of the economic system. In this category come the state of the weather, the development of foreign markets or of new resource bases in foreign countries, exogenous shifts in tastes and technology, and many others.

(ii) endogenous uncertainty -- uncertainty about the values of variables which will be determined within the domestic economic system once values of exogenous variables are known. Examples are domestic prices, tax rates, the level of demand, etc.

It is clear that indicative planning cannot act directly to reduce exogenous uncertainty -- though it is at least possible that by pooling and then disseminating the information available to all agents, each will be given a better basis for assessing these uncertainties. Because exogenous uncertainties will remain, an indicative plan must be a conditional or contingent plan: that is, it must take the form of a number of if ... then statements, each mapping out the course that the economy would take if some particular realization of the exogenous variables occurred. One might for example have one plan conditional on the OPEC cartel remaining effective, and another conditional on its failure.

Although indicative planning cannot reduce exogenous uncertainty and must consequently be conditional, it is certainly true that it may contribute substantially towards reducing the amount of endogenous uncertainty facing agents. If successful, it would provide them with a picture of the likely development of the domestic economy in each of a certain number of externally-generated contingencies, and would undoubtedly contribute to a reduction in the overall uncertainty facing them. One could regard this reduction in uncertainty/or, equivalently, the improvement in the quality of information available to decision-makers) as the main purpose of indicative planning.

Problems with Indicative Planning

Critics have pointed to several shortcomings with the concept of indicative planning. A point often made is that the repeated discussions of likely future developments that are an intrinsic part of the planning process may encourage the development of explicit or tacit collusion amongst the firms in an industry. Obviously, this is a potential problem, though probably not a particularly serious one: it is not clear that discussions within the framework of a planning apparatus would be more dangerous in this respect that those which normally occur in the context of trade associations and other industry organizations. Indeed there is a certain amount of evidence (see Estrin and Holmes) that in France, the country where this form of planning is most highly developed, planning has contributed to a reduction in the degree of cartelization.

A more serious problem is that the individual firms will have an incentive to distort the information they supply in order to create a more favorable environment in which to operate. This is a particularly accute problem where co-operative action is required, where the benefits of an individual's action are widely diffused, or where fiscal measures which will affect large numbers of firms have to be designed. For example, R & D may benefit many consumers and other producers who have an incentive to understate the value of these benefits in order to reduce the amount they might be asked to contribute, whilst the firm which is to be subsidized to undertake the research may overstate the costs in order to increase its profits. The implication of this is that the planning authority needs to be aware of any incentives that might exist to distort information, and to seek independent information if possible in such cases.

Another criticism often made, is that planning implies the development of a unified, economy-wide expectations about the future, in place of the multitude of possibly-divergent expectations one would expect to find in an unplanned situation. It is argued that this is harmful because if there are unresolvable uncertainties about the future, any risk-averse society should allow a dispersion of expectations rather than achieve a consensus because there is a risk that this concensus would be incorrect. In such a situation all agents in the economy would make bad decisions, whereas with a dispersion of expectations one would expect that in any given outcome only a limited number would do so.

It should in fact be clear that this criticism is based on a misinterpretation of the planning process. As described above, this process did not involve any attempt to reach a concensus expectation about exogenous uncertainties. It set out a number of possible paths for the economy, each contingent on a particular realization of the exogenous variables. Agents were completely free to assign their own probabilities to these various realizations. Planning involved establishing a concensus of expectations only about the behavior of endogenous variables, contingent on the exogenous ones.

A more substantial problem for successful indicative planning is likely to be posed by the fact that it would require a government to reveal its future fiscal and monetary intentions, and remain reasonably firmly committed to these. In the case of energy, for example, it would clearly not be possible to forecast even conditional developments without information on the aggregate demand management policies the government intends to pursue. Given the political complexities of the budgetary process, and the division of responsibilities between Congress and the Federal Reserve, such an unambiguous statement of intentions could be very difficult to obtain. Given its importance, it might however be worth lobbying for.

There are two more serious reasons why a government may not be able to, or may not wish to, reveal its future fiscal intentions. Newbery [1976] has shown that if a government cannot be contractually bound to its announced import tariff then it will wish to continually change the tariff. (The reason for this is discussed in the section on trade policy below.) Announcing its fiscal intentions will then reduce the credibility of the whole indicative plan as these intentions are changed.

Perhaps more fundamentally, there is an intrinsic contradiction between announcing a set of future tax proposals in a representative democracy such as the U.S. since subsequent administrations cannot be bound to them. The best that one can hope for is that a sufficient proportion of these plans are sufficiently widely supported to command continued support, or at least not to attract subsequent opposition. The problem of administrative credibility is sometimes advanced as a serious obstacle to long term decision making in the U.K., where successive governments have radically different policy objectives.

Conclusions

The potential contribution of indicative planning is difficult to evaluate without considerably more study at both the theoretical and empirical levels. It is clear that its contribution would lie in reducing endogenous uncertainties, and the potential significance of such a contribution would depend upon:

(i) the ratio of exogenous to endogenous uncertainties. In a situation where the principal uncertainties are all exogenous, planning would contribute little other than perhaps a pooling and dissemination of information.

(ii) the extent to which a planning procedure could be devised to cut through the complex interdependencies of a general equilibrium system and produce consistent and accurate forecasts.

As mentioned, it is difficult to be precise on these issues: there clearly are major exogenous uncertainties in the energy market -- the behavior of OPEC and the development of new technologies, for example. But it is also true that under rapidly changing market conditions, many of the uncertainties are endogenous, and stem from an inability to predict the reactions of other parties in the market under circumstances greatly different from any that have obtained in the past. It does seem likely that the institutionalized consultation involved in indicative planning could be particularly valuable under such conditions. The difficulties mentioned under (ii) above are purely technical in nature, though their magnitude should not be underestimated. However, the F.E.A. now has considerable expertise in the field of energy sector modelling and forecasting, and the technical requirements of good indicative planning are probably just within the state of the art.

3. Monopoly Policy

Theory of Imperfect Competition in the Supply of Exhaustible Resources

It may be argued that the extractive resource industries are more susceptible to cartelization because these resources are localized and immutable. The petroleum, tin, copper and bauxite cartels are examples of overt collusion in the supply of natural resources. However, the existence of a cartel is not synonymous with market power. Indeed, tacit collusion may be as successful in restraining competition as overt collusion, which implies that the absence of concentration cannot be relied upon as a reliable indicator of competition.

From conventional monopoly theory, we know that the market power of a monopolist depends on the elasticity of demand for the monopolized good and the ease of entry into the monopolized industry. Stated another way, if there are close substitutes for the monopolized commodity, an increase in the price of the good will cause consumers to shift to the substitute commodities. The limit on the monopoly price thus depends on the relative price of substitutes and the degree to which consumers are willing to accept other goods as a substitute for the monopolized commodity. Monopoly profits depend as well on the ease and speed with which entry into the industry can take place. The degree to which entry is successful in lowering the market price to the competitive level depends on the elasticity of demand facing each producer (see Chamberlain [1933]). For a homogeneous product like oil, entry can be expected to reduce the price (net of transport costs) to the competitive level.

One might argue that for nonrenewable resources, expansion of the industry in less than geological time is impossible. However, this ignores the fact while the resource base cannot be expanded over reasonable time periods, production capacity can be increased and known reserves can be enlarged through exploration. Entry may also take place through development of new sources of supply. Since the demand for natural resources is derived from the demand for final goods and services, it is not necessary that a substitute imitate the form of a particular resource, but rather it need serve some component of final demand previously served by the natural resource.

There are three ways in which the theory of monopoly for exhaustible resources differs from the theory for conventionally produced goods. The first has to do with rents. The first interim report described how the efficient price of an exhaustible resource may exceed the cost of producing the resource. The difference is a rent which accounts for the scarcity value of the resource. For a conventionally produced good, if the price exceeds the cost of production, one can assume that the market for the good is not perfectly competitive. This does not follow for an exhaustible resource, because even in a perfectly competitive economy the rent component of the price may be very large relative to the total price. Thus, for the case of oil, the fact that the OPEC price exceeds the cost of extraction is <u>not</u> sufficient evidence to assert that the OPEC price exceeds the competitive price. One might point to the correlation between the formation of OPEC and the rise in world oil prices as evidence for monopoly power. However, a counter argument is that the formation of OPEC coincided with a realization that the extent of world oil resources is a limiting factor in world economic development.[1/]

The second way that exhaustible resources differs from conventionally produced goods is the explicit interaction of the flow and asset markets for the exhaustible resource. A monopolist on production capacity (the flow market) may enjoy considerable short-run market power. However, if the resource base is not monopolized (the asset market), the expansion of capacity by competing owners of resource deposits will reduce the price to the competitive level. Conversely, a cartel may own sixty percent or more of the total resource base. Yet its monopoly power depends on its production capacity and the production capacity of competing suppliers. The determinants of monopoly power for an exhaustible resource is discussed in detail in the third section of both the first and the second interim reports. We will return to the distinction between monopoly power in flow and asset markets when we discuss policy alternatives below.

The third distinction that sets exhaustible resources apart from produced commodities is in the impact of demand elasticity on monopoly power. For a conventional static monopoly, the degree of market power (monopoly profits) is a decreasing function of the magnitude of the elasticity of demand. However, we have demonstrated in section 3 of the first interim report that if extraction costs are zero, and demand elasticity is a constant greater than one in magnitude, monopoly profits are identical to competitive rents. While there may be distributional issues associated with the ownership of the resource, the point is that for this case the price path given monopoly ownership is no different from the price path generated by perfect competition. However, if the demand elasticity falls below one in magnitude, the behavior of a monopolist departs radically from that characteristic of perfect competition. For this case, the owner of an exhaustible resource behaves much as if the resource were a conventionally produced good, and sets a limit price determined by competing sources of supply (see section 3 of the second interim report).

Policy Options

The major determinants of the market power of a resource cartel are its share of the total reserves of the exhaustible resource, its production capacity and the production capacity of competing suppliers, the price and availability of substitute sources of supply, and the level and elasticity of demand for the resource. Clearly these are not the only determinants of market power, but for analytical purposes they appear to be the major factors. For policy purposes, we may suppose that the share of total reserves (but not necessarily known reserves) and the production capacity of the cartel cannot be directly influenced by policy instruments. Therefore policy alternatives can be partitioned according to their impacts on the development of substitute sources of supply, the expansion of known reserves and production capacity outside the cartel, and the management of demand.

Both competitive rents and monopoly profits from the ownership of exhaustible resources depend on the difference between the extraction cost of the resource and the production cost of substitutes. Therefore, the development of cheaper substitutes for the resource will, all else equal, reduce the monopoly profits accruing to a resource cartel. Section 4 of the second interim report discussed in detail the optimal timing of research expenditures for the introduction of a substitute source of supply. The analysis in that section also demonstrated that a monopolist on a resource stock could be expected to underinvest in the development of a substitute for the exhaustible stock, and that the level of investment generated by a system of patent rights may not be optimal. The analysis in that section was not directly concerned with the distributional issues of monopoly ownership. Concern for the political power wielded by a resource cartel could justify accelerated development of substitute sources of supply.

It was shown in the second interim report that under some conditions, cartel profits are maximized by ignoring the exhaustible nature of the resource and simply maximizing current profits taking the production capacity of competitive suppliers as given. In this case, expansion of production capacity outside the cartel reduces the price of oil. Indeed, there is a multiplier effect. Each unit of capacity added outside the cartel forces the cartel to increase its own output in order to maximize profits. This suggests a basis for subsidization of domestic output capacity. Competitive suppliers of the resource consider the effect of capacity expansion on profits taking prices as given. However, each unit of additional domestic capacity increases the short run profit-maximizing cartel output and lowers the price. Since this is ignored by competitive producers, the rate of capacity expansion may be too low in the absence of a subsidy.

The importance of this finding is that the expansion of domestic production capacity could prove to be the most efficient means by which to reduce monopoly profits, less costly for example than the development of substitute sources of supply. Indeed, this seems to be the thrust of Project Independence. However, it must be explicitly recognized that an increase in current output reduces future production capacity. This implies lower output, higher prices and more dependence on imports in the future than would otherwise be the case. Since the government manages a significant fraction of the total resource stock of the country, it is important to determine the gains from an expansion of the rate of leasing of public lands. It is an empirical question to determine whether in fact cartel pricing policy is a function of competing production capacity, and whether an increase in domestic production is desirable.

A related issue is the expansion of exploration activities to increase the size of known reserves. For the purpose of dealing with a cartel, such a program is of secondary importance. The reason is that current cartel pricing policy should depend on the expected returns from exploration.

Demand management has a direct impact on monopoly profits. Conservation programs can be thought of as reducing the level of demand, while policies that increase flexibility can be thought of as increasing the elasticity of demand. Any increase in conservation lowers monopoly profits. An increase in demand elasticity may, in the long run, force the monopolist to increase the quantity supplied toward the efficient level (however, this does not eliminate scarcity rents). Also, if the increase in demand elasticity is anticipated by the monopolist, prices may rise in the short run. It is an empirical question whether the gains from such programs exceed the costs.

A closely related subject is the optimal tariff to reduce the level of demand and the level of dependence on imports. This is discussed in detail in the technical appendix and summarized in section 1.6 of this chapter.

4. Tax Policy and the Optimal Intertemporal Allocation of Oil

In this subsection, we review two sets of results: the intertemporal biases resulting from tax policy, and the use of tax policy to correct biases arising elsewhere.

In earlier reports, we suggested that there was an ambiguity in interpreting the bias associated with tax policy: most taxes introduce distortions. Were we to interpret a bias as a deviation of the tax induced policy from the optimal intertemporal allocation, or from the intertemporal allocation which would arise from some second best optimum (e.g. if all taxes were chosen optimally)? Alternatively, we could focus on differential treatment between the energy sector and other sectors, referring to any difference in the rates of return as a tax induced bias. Probably none of these are the appropriate measures for policy purposes: the appropriate question is simply whether any proposed change would increase welfare.

Several categories of distortions were analyzed: in the aggregate rate of consumption (use) of oil and in the pattern and technology of extraction and exploration, the tax provisions which were analyzed were the following:

1. Depletion Allowances. Our analysis suggested that the impact of a depletion allowance at a constant rate might be less than had previously been thought; if extraction costs were zero and the sector were competitive, it would have no effect. If extraction costs were positive and constant, it would lead to faster extraction (than in the absence of the depletion

allowance). If the sector were monopolized, it would also lead to a faster rate of extraction if the monopoly was excessively conservation minded (as it would be with positive extraction costs of a declining elasticity of demand), offsetting the effects of the monopoly; conversely if the monopoly was excessively profligate, the depletion allowance will lead to a slower rate of consumption.

The crucial parameters in determining the magnitude of the distortion for a competitive economy is the size of extraction costs relative to the price.

Although the effects of a constant depletion allowance may be smaller than is widely assumed, the effects of the gradual removal of the depletion allowance are larger than is widely assumed: if a depletion allowance is to be removed, it should be removed as quickly as possible.

Other provisions of the depletion allowance probably have serious implications for the industrial structure of the industry; what implications this has for economic efficiency is not clear. For instance, any provision which allows the depletion allowance for some categories of firms and not for others ought to result in most of the extraction being conducted by those who are allowed the depletion allowance. Provisions imposing a limit on the amount of the depletion allowance as a percentage of profits, unless computed on a well-by-well basis will also have an effect on the structure of the industry.

Although the depletion allowance would have no effect on the pattern of extraction from known reserves with zero extraction costs it does affect the incentive for discovering oil. The total stock of oil which is likely to be discovered is greater, and because of that present prices are lower, and consumption rates higher. Similarly, if there is a positive extraction cost, oil which it would never pay to extract without a depletion allowance will be extracted. But, apart from consideration of monopoly discussed elsewhere in this report, the cost of discovery and extracting this oil will exceed its benefits, and so such a policy is undesirable.

If the removal of the depletion allowance had not been anticipated, the immediate effect of the removal will be to raise the price of oil both because of the slower rate of extraction and the smaller stock of oil that will eventually be extracted. These effects may be disguised at present because the short run effect of a gradual removal of the depletion allowance is to increase the supply of oil.

2. Immediate Write-Off of Drilling Expenses. This provision provides the best example for the ambiguity associated with the meaning of tax induced bias. With immediate write off of capital expenditures, the corporation tax can be viewed as a pure profits taxes and is thus non-distortionary. But since other sectors are not so treated, there is a relative distortion: investment in oil exploration is encouraged relative to other kinds of investment. There is no justification for this favorable treatment.

3. Inability to Write-Off Immediately Expenditures on Lease (Or Land) Acquisition. For expenditures other than drilling, the oil sector is treated like any other sector. But this symmetric treatment does induce a distortion in the rate of extraction relative to what it would be in the absence of taxation. Consider a firm which purchases land under which there is oil. After the oil has been extracted, the land will be worth less.

The firm can take a capital loss on the reduction in land value, and thus reduce its tax liability. The present value of the tax write-off is increased by accelerating extraction. It would probably be desirable to allow immediate write-off of these expenditures; to discourage excessive allocation of resources to this sector relative to others from this favorable treatment, other taxes would have to be imposed.

4. Special Treatment of Capital Gains. Since the return to holding oil is the increase in its price, if this increase in price is subjected to capital gains taxation, the return to holding oil will exceed the return to holding conventional assets, and there will be excessive conservationism. This effect would be significant, since the difference in rates is large. Since the sale of oil is treated as ordinary income, for this distortion to occur requires that firms that do the exploration hold on to the land until the date of extraction. They then sell their land; the purchaser then extracts the oil, the income from the oil being perfectly offset by the reduction in the value of the land.

Of the various provisions discussed above, several have been introduced to offset the tax induced biases of the general tax structure (or so proponents of the provision argue). The depletion allowance is justified as a simple alternative to depreciation, but the present tax code essentially allows triple depletion: immediate write-off of drilling expenses, depletion allowances, and write-off of the loss in value of leases or land after extraction is completed. Other proponents of these provisions admit that it constitutes favorable treatment, but argue that the favorable treatment is required to offset the distortionary affects of the corporation tax structure, which unduly penalizes risky and capital intensive sectors. Elsewhere it has been argued that the sector is not more risky than other sectors, that the corporation tax does not discourage (but rather encourages) risk taking, that it does not penalize capital intensive industries, and that there are second best arguments for taxing the oil industry at a higher rate than other industries, because the share of rents in that sector is larger.

It is clear, however, that the instruments that we have analyzed, depletion allowances, special provisions for deductability of certain capital expenditures, and the special treatment of capital gains, may be used to offset distortions in the market allocation of oil, caused by one of the other market imperfections analyzed in other sections of this report.

5. Project Independence and Flexibility

The objective of national energy independence received widespread support following the Arab embargo; subsequent studies have established the high cost of total energy independence, and the goal has been consequently significantly modified. The question remains, however, of the economic rationale for independence: is operation independence only a political issue, or are there grounds for governmental intervention? And if so, what are the appropriate instruments for the government to use.

In a world of competitive markets in which individuals correctly foresee future prices, it is hard to make a compelling case for government intervention: if individual judgments concerning the probabilities of a future embargo (or future prices) are no worse than the collective judgments of government officials (and there is no reason to believe that they should be) then the actions they take, with respect to patterns of consumption of energy will be socially optimal; that is, for instance, if there is a finite probability of a very high price of oil, and the cost of constructing a convertible boiler is not too high, firms will do this, no additional incentives are required from the government.

The argument then for government intervention must be based on (a) the view that the market incorrectly assesses the probability distribution of future prices: (b) in contingencies such as the Arab boycott, prices are not allowed to rise to reflect social costs, so that firms and individuals have incorrect incentives for adopting technologies with sufficient flexibility; (c) part of the benefit of adopting more flexible technologies accrues to consumers in contingencies in which the price is lowered more than it otherwise would be (i.e. consumer surplus is increased); and since this is a social benefit not captured by firms, they will not adopt sufficiently flexible technologies.

In the appendix we develop a measure of flexibility, establish the intertemporal tradeoff between flexibility now and flexibility in the future, and derive conditions for the optimal intertemporal distribution of flexibility; this entails, with constant elasticity demand functions and approximately zero extraction costs, that the ratio of imports to consumption remain constant.

If firms incorrectly perceive the probability of an embargo, the optimal intertemporal allocation of oil may be enforced by percentage quota. (A tariff will not be able to enforce the correct intertemporal allocation; if, for instance, firms assumed the probability of an embargo is zero, a small tariff will leave unaffected the intertemporal allocation of the extraction of oil, except in the case where firms were exactly indifferent as to when they extracted the oil.)

If, on the other hand, the probability of an embargo is correctly perceived, but the extraction of oil is monopolized, then a tariff or subsidy can be used to correct the allocation of oil; although our analysis is able to identify the parameters which determined whether the monopolist would extract too fast or too slowly for optimal intertemporal independence, there were off-setting effects, so that without detailed quantitative analysis, it was not possible to say whether a tax or subsidy was called for.

If firms assume that the probability of a price rise by oil exporters is zero, then, if there are costs of adjustment (associated with the consumption of oil, e.g. of the capital stock associated with it), the market will lead to excessive consumption of oil. This can be corrected again by the imposition of a tariff, the rate of which is equal to the expected percentage rate of price increase.

6. Trade Policy

In part III we examine the problem of setting taxes on the production and import of exhaustible resources. A pure rent tax on producers is non-distortionary and we show that an appropriately designed ad valorem tax can, under certain conditions, produce exactly the same non-distortionary effect as a rent tax when levied on competitive domestic producers. This suggests that for a large importing country like the U.S. an ad valorem import duty may be optimal, and indeed this would be the case if the importing country could be contractually bound to its announced tax schedule. However, in a world of competing nation states, such contracts cannot be enforced, and consequently have no credibility. In such a case there will be no optimal tariff on competitively supplied exhaustible resources, even though the optimal tariff on produced goods is well defined.

The reason is simple. Foreign suppliers of oil, say, have to decide whether to sell today or leave it in the ground for future sale. If the future price cannot exceed the cost of producing a substitute from the "backstop technology," then the present price is formed by working back from the future price, ensuring that supplies are just exhausted when the backstop is introduced. An importing country will choose a tariff which reduces its dependence on foreign oil and thus reduces world demand. The price will fall to induce the same trade to be demanded, so that over the life of the oil fields the same total is extracted. The benefit to the importing country of the tariff (or reduced demand) is reflected in the lower price. The tariff must rise at the rate of interest to make the importing country indifferent to when it imports oil at the margin. It follows that the future tariff is higher than the present tariff, imports will be gradually reduced, and all consuming countries will benefit from the currently lowered price of oil.

However, having imported the currently cheap oil (cheap because of the announced high future tariff, and low future demand) the importer would like to revise his tariff so as to consume more oil later. It was, after all, only the implicit promise to import less in the future which lowered the price, not the act of so importing. Since the promise cannot be enforced, and since the producing countries know that it cannot, the suppliers have no rational basis on which to determine their current supply schedule. To be precise, if the large importer takes the behavior of the other importing countries as given, and if foreign oil producers are competitive, and if they forecast future prices by assuming that the importing countries behave rationally, then there is no optimum tariff.

We show that this paradox can be partially resolved if suppliers are myopic, or if the government does not alter its tariff (which would not be in its best interest, but may be a good description of likely administrative behavior), or if there is sufficient uncertainty about the future price of oil.

7. The Rationale for Price Control Programs

The classical arguments against price controls rest on the assumption that markets function efficiently and income can be redistributed without distorting market incentives. Interfering with the market-clearing price makes necessary quantity rationing on either the demand or the supply side of the market. Unless the rationing authority knows each consumer's valuation of the resource, the resulting allocation will be inefficient in that some consumers would be willing to buy from others who would be willing to sell, and a black market may develop. Allocation schemes that permit trading (a white market) may preserve efficiency.

Markets for exhaustible resources clearly are not perfect. The first interim report detailed the possible instances of market failure in the long-run allocation of exhaustible energy resources. Of particular concern for the question of price regulation are imperfect competition, absence of complete markets and informational externalities.

It may be argued that all practical systems of income redistribution affect production and consumption decisions over time, and hence non-distortionary income transfer is not feasible. Given this constraint on income distribution, there may be a basis for price controls on energy resources. The income elasticity of demand for energy resources is relatively low and the owners of energy resources are certainly not the poor. Therefore, energy price increases have a larger negative impact on the welfare of the relatively poor than the relatively rich. Price controls and rationing supposedly ameliorate this undesirable impact on the distribution of income.

However, the net benefit of price controls to protect income equality depends on the effects of controls on production and consumption decisions. This remains an area for further theoretical and empirical research.

The control of market prices, such as the wellhead price of natural gas, may be viewed as a subset of a wider class of price controls. For example, rather than specifying the wellhead price of natural gas, the regulatory authority could set a lower and upper bound on the price. The lower bound would be a price guarantee to the producer, much the same as the support price for certain agricultural products. Alternatively, the regulatory authority could be more specific and set prices corresponding to particular events that affect the supply and the demand for natural gas.

There is an important distinction between regulated prices and price subsidies or taxes. A tax or subsidy is equivalent to a shift in the supply or demand of the commodity affects. These controls change the market clearing price, but the extent of the change depends on prevailing demand and supply conditions. Regulated prices, on the other hand, peg the market price at a particular value, or constrain the price to lie within a particular range. The difference is illustrated in figures 1(a) and 1(b). The effect of a tax of magnitude τ per unit of the commodity levied on the producer is shown as a shift in the supply curve. The equilibrium after tax price, p_1, depends on demand and supply elasticities. In a competitive market if demand is inelastic relative to supply, the tax will increase the consumer price, while if supply is inelastic relative to demand, the tax will have less impact on price but will reduce the amount of the commodity supplied. The effect of a regulated price is very different. If the regulated price equals the market equilibrium price (or if the market equilibrium price lies within the upper and lower bounds of the regulated price), then clearly there the price regulation has no effect, and there is no need for regulation. If the regulated price is below the market-clearing price (or if the market price lies above the upper bound of the regulated price), then the regulating authority must ration demand for the commodity. On the other hand, if the regulated price is above the market price, (or if the market price lies below the lower bound of the regulated price) the authority must purchase the commodity to maintain the administered price. This is illustrated in figure 1.7b where Q_r is the amount that must be purchased by the regulatory authority to maintain the price p_1.

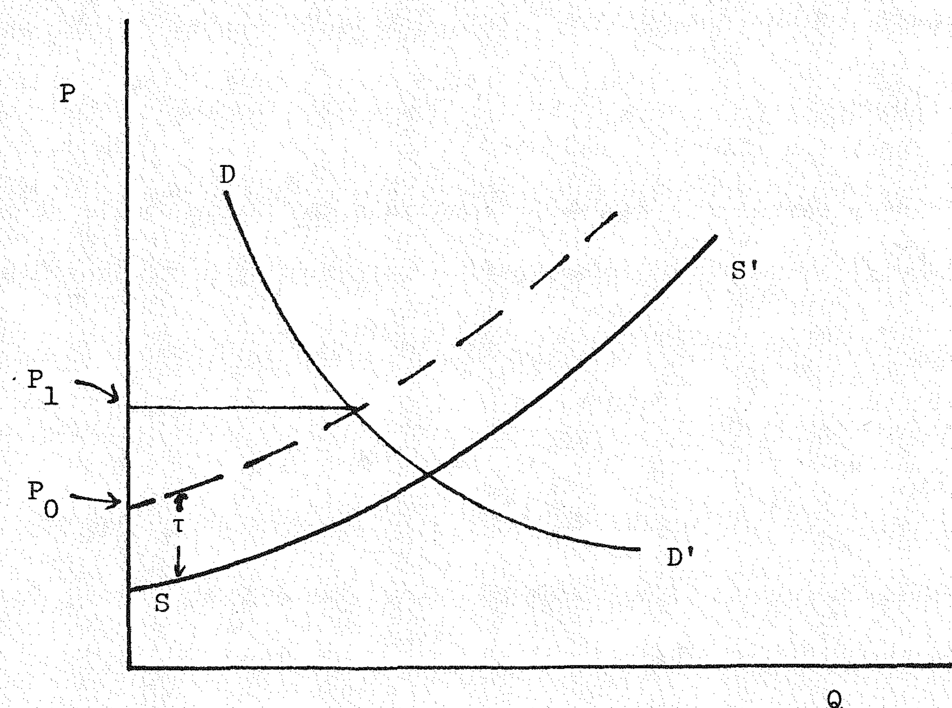


Figure 1.7a

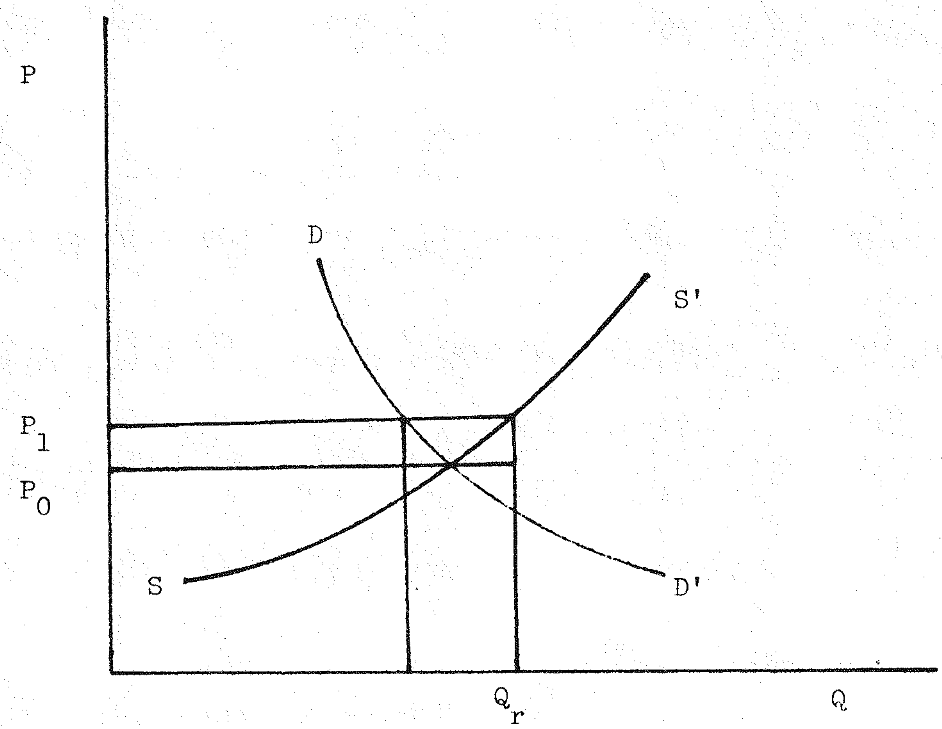


Figure 1.7b

While taxes and subsidies alter supply or demand, price regulation impairs the allocating function of the marketplace. If the price constraint is active, the market-clearing price is not directly observable, and hence the extent of the distortion imposed by the regulation is not directly determinable. Taxes and subsidies are often the preferred mechanism to correct for costs or benefits that are not properly accounted for by the market. However, we will demonstrate that in cases in which the market for a particular good or service is absent, the imposition or a regulated or administered price may be justified on the basis of economic efficiency.

Indeed, in these cases price regulation acts to create a market where no market would exist without intervention.

The advantages and disadvantages of price regulation to limit undesirable distributional effects have already been mentioned. There are at least three other motivations for price support programs. These are stabilization policy, support of research and development, and risk-sharing. In each of these areas there is a common thread that weaves the basis for a price support or guarantee programs, namely the absence of a complete set of markets to allocate the goods and services of value in the economy.

Stabilization Policy

It is useful to distinguish long-term from short-term stabilization goals. Long-term stabilization policy is concerned with eventual surpluses or deficits in resource stocks. Economic efficiency calls for an "orderly" transition to substitutes for exhaustible resource stocks (in the sense described in detail in the second interim report). If the resource is consumed too quickly, future prices must rise faster than the rate of interest, and if the resource is consumed too slowly, surpluses will persist indefinitely. Long-term stabilization requires that the current resource price be consistent with the optimal use of the resource over all time.

It is possible to stabilize at the long-term utilization of an exhaustible resource through price regulation, however such a policy is inferior to alternative measures. The difficulties are easily illustrated. If the rate of utilization is too rapid, the regulatory authority would have to purchase and store a continuously accumulating stock of the resource in order to raise the price to the optimal level. The extraction and consequent storage of the resource incurs social costs that reduce the benefits of the stabilization program. If the rate of utilization were too conservative, the government would have to supply the resource, presumably from its own stockpile, in order to reduce the price to the socially optimal level. A further important consideration is that the optimal price, and hence the optimal price support program, depends on demand and supply conditions and will be a complicated function of time. If this is not recognized, the support program would exacerbate the intertemporal resource allocation.

There are two fundamental reasons why long-term stabilization measures may be necessary. The first is the absence of risk and forward markets extending indefinitely far into the future. Without these markets, it is impossible for individuals to observe directly through forward transactions whether current price levels are consistent with long-term efficiency. In other words, there are no direct market signals for whether resource utilization is too profligate or too conservative. If these markets did exist, and resource consumption were too profligate, there would be excess demand for the resource at some future date and profitable incentives to correct this excess demand.[2/]

Long-term stabilization measures may also be necessary to correct for distortions in the market rate of interest. It is clear from the analysis presented earlier in this program that the optimal allocation of exhaustible resources depends critically on the social rate of discount. Since this may not be equal to the private discount rate, there may be a basis for long-term measures to change the market allocation or resources.

The correction of the private rate of discount can be influenced through monetary policy. The importance of forward markets is primarily an information problem. If demand and supply constraints were known, the appropriateness of current rates of utilization could be determined. This would mitigate the need for forward markets.[3/] Hence the regulatory authority could promote long-run stability by collecting and disseminating information about future economic conditions.

Even if, on the average, the long-run allocation of resources is optimal, there may still be problems associated with short-run instabilities. Short-run instabilities may be a particular problem for a resource like oil, for which there exist simultaneously both an asset and a flow market. The problem is that a speculative rise in the asset value of oil may prompt producers to reduce supply, and hence the price of oil may rise, reinforcing the original speculative price rise. The end comes when traders realize that the rate of price rise cannot be sustained, but until then there are real profits to be made by riding the speculative wave.

With a complete set of futures markets, the instability induced by speculation cannot occur, since trade in forward markets insures that the rate of price change cannot exceed the equilibrium rate of return on assets.[4/] Thus we see that once again the underlying difficulty is the incompleteness of markets.

Short-run instabilities can be removed or ameliorated through price regulation. In the latter case, the range of admissible price variation may be specified by regulating the upper and lower bounds on price. This price regulation could be implemented by holding buffer stocks that are depleted when price exceeds the regulated upper bound, and increased when price falls to the regulated lower bound. This is the general policy employed in agricultural price support programs and foreign exchange stabilization programs.

A short-run stabilization program is ideally suited to cope with the risk of embargo. The buffer stocks may be held in above or below ground storage, or shut-in capacity. In the absence of embargo, the maintainance of a buffer stock of oil increases the social cost of oil by some amount. This is characteristic of all price support programs, since stabilization does not come at zero cost. The gain, clearly is a reduced price in the event of an embargo. This can be thought of as short-run rather than long-run stabilization since the program is not intended to change the long-run allocation of oil.

There are two important considerations in a price support program. First is the technical question of how to implement the stabilization program. This includes a host of important factors. The admissible price variation must be specified. The decision must be made to ration supply or supplement supply with the use of buffer stocks, or some combination of the two. The buffer stock program, if one is used, should be designed to achieve lowest costs. For example, for stabilization against foreign oil supply interruptions, the optimal mix of storage and shut-in capacity must be determined.

The second major issue is the separation of short-run from long-run stabilization policy. Short-run price control is feasible because by definition it is not necessary to provide for a continual drain on resources of the regulatory authority. For example, the long-run cost of maintaining a buffer stock must be paid, but it is not necessary to continually accumulate or deplete stocks for short-run stabilization. Unfortunately, it is difficult to discern a long-run stabilizating trend from a string of random occurrences with zero mean. In other words, it is easy to rationalize a long-run trend toward higher oil prices with the argument that exploration efforts have been unlucky, and a large discovery is around the corner. In the case of exhaustible resources, one expects price variation (in real terms) in the long-run, and hence short-run price policy must be adjusted to conform to the long-run trend line. If this is not done, the regulatory authority places a net subsidy or net tax on the resource sector, and thus performs a role very different from that of short-run price stabilization.

We have assumed that the elimination of price fluctuations is desirable, however, this need not be true in all cases. Clearly, there is a difference between price instabilities and price movements necessary to equilibrate markets with changing supply and demand. It has been assumed that stabilization policy reduces instabilities without interfering with equilibrating price movements. In practice, any stabilization program may have deleterious effects on incentives to provide information about future demand and supply, and on the ability of markets to reach equilibrium. This trade-off should be explicitly recognized in the design of stabilization programs.

Another caveat relates to conditions under which it may not be desirable to eliminate price instabilities. In the second interim report, we demonstrated that under certain conditions it would pay a monopoly supplier to randomize prices. Relative to a constant monopoly price, such a strategy need not lower total consumer welfare. Thus, price stabilization may not be desirable if the consumer is large relative to the total market.[5/]

Imperfect Competition

The second source of market failure that may motivate price regulation is the cartelization of resource supply. Of course price control is common in the regulation of franchised monopolies, and much has been written on the costs and benefits of utility regulation. We consider here the merits of price regulation as a defense against the monopoly power of a cartel that is not under the legal jurisdiction of the consuming nation. We have already discussed one special case of price regulation as a defense against monopoly power, namely the maintenance of buffer stocks in the event of an export embargo. Clearly this is a case of monopoly power on the part of the exporting country.

The first and second interim reports in this program analyzed in great detail the determinants of monopoly power for a resource cartel such as OPEC. (See section 3 of the first and second interim reports.) Under conditions of inelastic demand, the cartel price was limited by the production capacity of producers outside the fringe and by the production cost of an eventual substitute for the exhaustible resource. It was emphasized that under conditions of inelastic demand, patent rights may be ineffective in providing incentives for the development of substitute sources of supply. The reason is that the cartel will find it profitable to set the price of oil just below the price of a substitute source of supply, and if the resource stock owned by the cartel is large, the present value revenues from an alternative energy source may be negligibly small. Of course in a Utopian economy, it may be desirable to delay the introduction of substitutes until cheaper sources of supply are exhausted. However, if the cartel succeeds in maintaining a resource price above the optimal price level, then the introduction of substitute sources is delayed beyond the optimal date and the revenues that may be generated by an alternative energy source are correspondingly smaller.[6/] In an extreme case, the cartel may own enough of the resource to effectively eliminate private incentives for the development of alternative sources of energy. There are obvious gains to consumers of the resource in pursuing policies that force the cartel to lower its price (provided that the cartel price is indeed above the optimal price). This can be done by subsidizing the development of substitute sources of supply.[7/] The social gain from such a program is the resulting decrease in the cost of energy. It is not the promotion of research and development (neglecting unexpected benefits that may be derived from the new technology). It is plain that any price above the extraction cost of oil can be undercut by the OPEC cartel. If a substitute for oil (e.g. shale oil of coal liquefaction) could be produced on a large scale relative to world demand at $10 per barrel, $2 below the current OPEC price, OPEC could reduce the price of oil to $9.99 if necessary to preserve a market for its product. By assumption, this would eliminate the market for the new energy source. If the substitute were produced privately, the operation would be a financial disaster, even if the firm had patent rights on its product. From a social point of view, the operation was definitely not a disaster, since it succeeded in lowering the cost of oil by $2 per barrel.[8/]

The existence of competitive suppliers of oil, outside the OPEC cartel, restrains the monopoly power of the cartel. It is shown in the second interim report that a competitive fringe will limit the resource price so that it approaches the cost of a substitute source of supply over time. When the price of oil is low, so too are the benefits of promoting the development of alternative energy sources. Furthermore, the analysis revealed the importance of the fringe capacity on the market price of oil. It is the prospect of high future prices that provides private incentives for the fringe to expand production capacity; and it is the expansion of total fringe capacity that limits the monopoly price. Thus any social program to promote substitute energy sources must evaluate time path of benefits from research and evelopment. In particular, a laissez-faire policy wherein the competitive producers of oil solve the OPEC problem by expanding total production could be less costly than an active energy research program. This is a question that demands empirical study.

Assuming the judgement has been made to promote the development of alternative sources of energy, there remains the choice of the organization of such a program. The time path of investment must be decided, as well as the choice of technology. Such a program could be socially managed, supported entirely by government revenues. It is of more than academic interest to inquire if a system of incentives can be constructed whereby such a program could be efficiently decentralized. We have already argued that a system of patent rights is ineffective in this case. The ownership right on an invention is of no value if the market for the invention can be removed through price competition by the OPEC cartel. The problem is stated in this way to emphasize that patents fail because the market for the patented product does not exist. This suggests that direct price control may be an effective regulatory tool.

Let us return to the example where a substitute for oil could be produced at $10 per barrel, $2 per barrel below the OPEC price. The regulatory authority could guarantee a future price of oil of $10 per barrel. That is, it announces that it will purchase oil from anyone at $10 per barrel. Of course the contract would be limited to domestic producers, and conditions of the contract would be specified to distribute risk between the government and the producers in an agreeable manner. This contract is actually a revenue guarantee to producers, but it can be properly considered as a particular example of price regulation. Given current world oil prices, no one would volunteer to enter in this contract. However, if any substitute could be produced for the price of $10 per barrell or less, a market for the product is guaranteed.

Of course uncertainty plays a crucial part in this game. Suppose an enormous oil field is discovered, (or there is an unexpected technical breakthrough), and the world price of oil plummets below $10 per barrel. The government is committed to a contract at a cost in excess of the world price of oil. The original intent of the contract was to lower the price of oil. If the price falls for some other reason, the contract has negative social value, but this cannot be avoided ex ante. Another problem is that the cost of producing the substitute may have been over-estimated, perhaps intentionally by potential producers. A possible recourse is a clause to the effect that the contract will be terminated if the market price of oil falls below a specified value for a specified time. An alternative is the establishment of a franchised monopoly on the substitute product, the price of which would be regulated to permit a fair rate of return to the firm. As in the guaranteed price scheme, there is still the potential for government subsidy -- witness the case of U.S. railroad passenger service or the Postal Service.

Risk

The future value of an exhaustible resource depends on events that cannot be predicted with certainty. This implies that the decision to leave a barrel of oil in the ground is an investment decision under certainty. The return to this investment depends on such factors as the amount of new discoveries, technical developments and government regulations. Assuming producers of the resource are risk-averse and the market does not provide the opportunity to completely diversify investments (because markets are not complete and the set of feasible trades does not cover all states of nature that affect the value of the resource), there may be a market bias in the utilization of the resource relative to the socially optimal rate. A detailed analysis of the direction and the extent of the bias is presented in section 3.4 of the technical appendix of this report.

To substitute for the absence of forward markets, a regulatory authority may offer price guarantees to the owners of the resource. One possible policy consists of the government guaranteeing the resource owners the expected price at some future date, thereby removing the uncertainty facing resource owners. Alternatively, and more in keeping with policies adopted in agricultural commodity markets, the government might guarantee not an exact price, but a lower limit to the price.

The direct control of price may be used to serve several possibly conflicting goals. However, the objectives of stabilization, support of research and development and the sharing of risk are consistent in that in each case the need for price control stems from the absence of a complete set of markets for the resource. Thus a price control program designed to serve these goals would operate most efficiently if the controls modelled the behavior of the resource price that would be expected if markets for the resource were complete.

8. Directions for Future Research

Although we are pleased with the progress made under this contract there remain a number important directions for future research. First, and perhaps most important, is the actual utilization of our parameterizations. This would require some detailed econometric analysis, and although Microeconomic Associates feels hesitant at the present time to undertake this task itself, it would be willing to work closely in conjunction with either the FEA itself or another contractor in the empirical implementation and further development of the parameterizations presented in our three reports.

In our judgment, the work on the effects of imperfect competition is the most ready for immediate implementation and needs the least further development.

In all the other areas noted, although we have made significant progress, further work seems called for. The most promising areas, combining a high likelihood of significant analytic advances with important policy implications, are the following:

(1) Uncertainty: The analysis in the Technical Appendix establishes the parameters which determine the magnitude and direction of the bias in extraction arising out of the absence of a complete set of risk markets. However, we have been unable to establish any presumption for under or over extraction and further work is necessary. In addition an important set of questions concerns the value of obtaining better information to reduce uncertainty about the potential supply of natural resources and substitutes, about market demand and about other environmental uncertainties. Although in the present report we discuss the role for government collection of information through indicative planning, much further work remains.

(2) Taxation: The analysis of the distortionary effects of taxation was more complicated than expected and yielded results (e.g. relating to the effects of depletion allowances) that were not entirely expected. Again, although significant results were obtained, our analysis leaves us convinced that further work is called for. For instance, tax policy may effect not only the rate of aggregate extraction but also the rate of extraction from any particular well. This is important, because it will effect the aggregate supply of oil that can be extracted. An analysis of this will require more detailed modeling of the process of extraction than we have been able to undertake within the scope of the project.

(3) Flexibility: Our analysis focused on the implications of the possibility of an embargo (or a sudden price rise) for the intertemporal allocation of a natural resource. The possibility of embargoes also has important implications for the choice of modes of consumption and production, i.e. of the elasticity of the demand and supply functions. The analysis undertaken in conjunction with the second interim report and this report convinces us that there may be systematic biases in the choice of "flexibility" the exact nature of which would require considerable further analysis.

(4) Regulation: The project was concerned with intertemporal inefficiencies in resource allocation, not static inefficiency. Yet there are important interactions between the two requiring detailed analysis of the static inefficiency. In particular, we were forced to ignore one important institutional aspect affecting the utilization of energy resources in the United States: that much of the oil and gas is

consumed by electricity generation which is a regulated sector. Thus, this portion of the demand for oil and gas is a derived demand, and a derived demand not based on competitive markets. What implications this has remains an open question.

A second aspect of regulation relates to incentives, e.g. for the development of lower energy consuming automobiles or expanding the supply of oil. Conventional economic analysis argued that prices provide the best incentives. Yet under a variety of relevant "second best" circumstances it can be shown that regulation, or at least non-linear pricing, is preferable to pure price system. Do these circumstances obtain in the energy market? If so, what kinds of regulations ought to be introduced?

A similar question arises with respect to rationing in the event of a shortage. As we discuss in this report, such price controls are likely to have significant deleterious effects. Can these be mitigated somewhat by the careful design of the rationing scheme, and if so, how? For instance, is a "white market" desirable? It is important that such a scheme be worked out in detail prior to the imposition of the embargo.

(5) Stabilization: An alternative approach to the threat of unstable foreign prices which is an alternative to flexibility and independence (or perhaps complementary to it) is the utilization of domestic buffer stocks either below ground or above. For policy purposes, it is important to know how best to design such a buffer stock scheme, what rules it should follow. Moreover, there is a trade-off in the degree of flexibility in, say, demand and the required (or optimal) size of the buffer stock. There are reasons to believe that the market will not

make the correct decision in this respect. The appropriate government policies to counteract these deficiencies need to be developed.

(6) Technical Change: Again, although we feel that significant progress has been made in the analysis of the market incentives for the development of substitute sources of energy, much remains to be done. The models we have employed have not treated uncertainty associated with research fully. They have focused on one dimension of the decision, the level of investment, and not on the choice of research projects (e.g. whether overly risky research projects are undertaken under particular market structures).

(7) Cartels: As we noted earlier, perhaps the most significant progress made in the course of this contract was the analysis of the behavior of a resource cartel. However, as our general understanding of resource markets increases, we will need to extend our analysis of cartel behavior, e.g. to account for the reaction of the cartel in the effects of policies on the part of the importing nation.

(8) Social Discount Rate: The intertemporal allocation of exhaustible resources depends crucially on the rate of discount (which, in efficient equilibrium is the social discount rate). The actual discount rate will differ because of taxes, imperfect capital markets (which are particularly acute for international capital transactions) and problems of moral hazard and bankruptcy. The relationships between the efficient and market interest rate (or, more properly, rates) is complex, and although much research has been directed to this relationship, our understanding remains seriously incomplete.

The list is a long one. It is conventionally said that all good research opens up more questions than it closes. While this has been our experience in the course of this project, we feel the results underscore the applicability of economic modeling in this policy area.

II. Review of the Literature on Market Biases in the Intertemporal Allocation of Energy and Other Natural Resources

1. Economic Efficiency and the Intertemporal Allocation of Resources

1.1 Introduction

The standard framework for analysing market biases in the intertemporal allocation of exhaustible resources is first to analyze the optimal intertemporal allocation of resources, then describe the market allocation of resources, and finally to compare the two. Although this framework seems perfectly reasonable, there are a number of problems associated with its implementation, which we address in the subsequent subsections.

Among these problems, perhaps the two most important are:

1. What is the appropriate criterion for optimality?
2. What is the appropriate description of the market economy?

The first is concerned with the standard questions of welfare economics, and the section below is concerned with applying the "conventional wisdom" on these questions to the particular questions associated with the depletion of natural resources.

Conventional welfare economics addresses the allocative efficiency of a decentralized price system, and the role of government intervention to correct instances of market failure and to achieve certain distributional objectives. The need to distinguish between an efficient and a socially optimal allocation of resources is clear. A perfectly competitive, static economy with no uncertainty will sustain an efficient allocation in that no individual's welfare can be increased without decreasing another's. Such a distribution of the products of the economy may, however, be very unequal. Optimality is concerned then also with the distribution of the outputs of the economy to all members of the society. Normally, efficiency is considered a necessary, but not sufficient condition for optimality.

The central feature of exhaustible resources is that their current use denies their primary services to future generations. Thus an economic analysis of exhaustible resources must explicitly consider the allocation of resources over time.

1.2 Analogies between Intertemporal Allocation of Resources and Static Allocation

It will be convenient if we divide the literature on intertemporal resource allocation into that which is concerned with the basic similarities between the problems of intertemporal resource allocation and static resource allocation (section 2), and that which is concerned with the important differences between the two (section 3). Although much of the literature to which we refer in the discussion below is not immediately concerned with the intertemporal allocation of natural resources, this literature does provide the necessary framework into which such a discussion must be placed. Although our discussion is divided into these two parts mainly on grounds of analytical and expository convenience, the relative importance of the similarities on the one hand and the differences on the other between intertemporal and static resource markets has important policy implications, and there is not universal agreement about whether, say, the kinds of remarks made in section 3 should be treated as footnotes to the basic discussion of section 2, or whether section 2 should be treated as an interesting but not very essential prelude to the core discussion of section 3.

1.2.1 Efficiency

Malinvaud [1953] demonstrated that, under certain assumptions, it is straightforward to extend the precepts of static welfare economics to an intertemporal setting.[1] Since two units of a commodity delivered at a particular time is different from one unit of the commodity delivered at two different times, the analysis must differentiate inputs and outputs according to the date at which they are made available. A similar construct was advanced by Arrow [1964] and Debreu [1959] to extend the concept of decentralized efficiency to the uncertain environments which naturally arise in an intertemporal economy. Here it is necessary to distinguish commodities according to the state of the world in which they are made available.

1.2.2 Distribution

Similarly, there is a close analogy between the problems of intertemporal distribution of income and those of the distribution of income at any point of time. Just as there is no presumption in a static model that the distribution of income among individuals would be socially desirable, so too there is no presumption in an intertemporal model that the distribution of income among generations would be socially desirable.

The determination of whether a particular allocation is or is not socially desirable depends on two important sets of assumptions: those concerning the structure of the social welfare function, and those

---

[1] Some additional conditions arising out of the infinite horizon must be satisfied. See Malinvaud [1953] and below.

concerning the mechanisms by which the market allocation may be altered. In a static environment, the former is concerned with the strength of the equalitarian ethic. At the one extreme, Rawls [1971] has recently suggested the strongly equalitarian objective of maximizing the utility of the worst off individual. This objective function still, of course, does not resolve the classical problem of interpersonal comparisons; that is, to ascertain who is the worst off individual requires strong interpersonal comparability. The criterion has been used not only in the static literature on income redistribution (Phelps [1975], Atkinson [1975]) but has also been applied to the problems of intertemporal allocation of natural resources (Solow [1974]). However, there have been many criticisms of the Rawlsian formulation, some viewing it only as a limiting case of the more conventional utilitarian formulation (Arrow [1973], Harsanyi [1975]), others focusing on the particular problems of applying the Rawlsian formulation to intertemporal problems (Dasgupta [1974], Arrow [1973]).

Within the utilitarian framework, a constant elasticity of marginal utility formulation,[1]

---

[1]Although the constant elasticity of marginal utility formulation is a reasonable, easily implementable parameterization, there is not universal agreement to it. It implies that a 1% increase in income reduces the marginal utility of income by the same amount, regardless of the level of income. Kolm [1972] has criticized this assumption, and suggested that a constant absolute risk aversion utility function (an exponential utility function) is a more reasonable parameterization. Both of these parameterizations have strong implications when applied to risk analysis, and in that context probably neither are completely acceptable. (See Cass-Stiglitz [1970], Stiglitz [1972]). (On the appropriateness of using parameterizations developed for risk analysis to problems of distribution see Harsanyi [1955]).

$$u = c^{-\alpha+1}$$

(where u is utility, c is (lifetime) consumption, α is the elasticity of marginal utility) has been found to be a simple, easily implementable parameterization. It is sometimes argued that a reasonable value of the elasticity of marginal utility -- based on the analogy between attitudes towards risk and attitudes towards inequality -- is between 1 and 2.[1]

Atkinson [1970] has interpreted the elasticity of marginal utility as a measure of inequality aversion. For small variances in the distribution of income, society would be willing to give us a percentage of national income equal to 1/2 times the coefficient of variation squared times the measure of inequality aversion to be rid of all inequality. Stiglitz [1975a] has developed a corresponding marginal measure of inequality aversion; if 1% of the income of each individual were taken away, and an identical absolute amount given to each individual, in the process of redistribution, what percentage of the revenue collected could be dissipated away and still leave society as well off as before the redistribution? He has shown that for small variances the marginal measure is twice the Atkinson total measure.[2]

The second important parameter in the intertemporal social welfare function is the pure rate of time discount. Although there may be some evidence that individuals have a positive pure rate of time discount,[3]

---

1 Other approaches to the measure of the elasticity of marginal utility are discussed in Fellner [1967]. The early development of this approach is due to Fisher and Frisch.

2 These parameterizations focus on the intertemporal distribution of aggregate consumption. Just as it is important to include differences in leisure in the measurement of the static distribution of income, so too here. See Stiglitz [1975a].

3 Indeed, this provided one of Bohm-Bawerk's three reasons for a positive interest rate.

there is no reason that society ought to have any positive discount rate (as Ramsey argued [1928]). On the other hand, if there is uncertainty about the life of the world, then under certain conditions it can be shown that the maximization of expected social welfare can be described simply by imposing a discount rate (as Yaari showed for uncertainties of individuals [1965]). Dasgupta [1973] and Dasgupta and Heal [1974] have recently applied a similar formulation to a problem arising in exhaustible natural resources (to be described in greater detail below); in their formulation, it appears as if the conditions under which simply adding a discount factor is appropriate are very restrictive.

The importance of the precise number assigned to the elasticity of marginal utility and the pure rate of time discount can be seen easily as follows:[1]

Assume that income will be rising at 3% a year over the next 25 years. Then how much society would be willing to give up today, to obtain a dollar in 25 years is given in Table 1.

---

[1]Throughout this discussion, we have assumed an additive social welfare function. Koopmans [1960] has criticized this formulation, and proposed a more general stationary utility function. However, no easily implementable parameterization of this class of functions has been developed.

Table 1: Amount society would trade today for one dollar in 25 years, assuming income is rising at 3 percent per year.

$$r = \delta + \alpha \frac{\dot{y}}{y}$$

r = consumption rate of discount

δ = pure rate of time discount

α = elasticity of marginal utility

y = income

		α = 1	α = 2
	0	$.47	$.22
δ	.05	.14	.06
	.10	.04	.02

TABLE 1

1.2.3 Mechanisms for Altering the Allocation of Resources in a Mixed Economy: the Second Best Problem

The importance of the second set of assumptions, that concerning the mechanisms by which society can alter the allocation of resources, has been widely discussed in the static literature, but the corresponding questions have not been extensively discussed in the dynamic literature. In the static literature, for instance, the distinction between those allocations which can only be obtained by lump sum redistributions, those that can be obtained by some income tax structure, and those that can be obtained by some linear income tax structure is well known. Much of the older literature in static welfare economics had a very unreal quality to it, because it assumed that all lump sum redistributions were feasible. Allocations which are obtained by lump sum redistributions are sometimes referred to as first best optimum, while those in which such redistributions are not feasible are sometimes referred to as second (or third) best optimum (depending on what additional constraints are implied).

The major instrument for the intertemporal redistribution of income, at least as it has been viewed in the recent literature, is monetary policy. Under certain reasonable conditions, it can be shown that monetary policy can attain any (or at least any within a wide range) of intertemporal redistributions. If this is the case, then an analysis of the intertemporal allocation of exhaustible resources need only focus on efficiency questions, and not on distributional issues, leaving those for monetary policy.

Although it may be _feasible_ for monetary policy to obtain any desired intertemporal redistribution, the monetary authorities may not be centrally concerned with this issue, but rather focus their attention on short run issues of economic stabilization. In that case, it does appear necessary for the policy concerning intertemporal allocation of resources to concern itself with intertemporal distributional issues.

In short, the issues of whether natural resource policy ought only to be concerned with intertemporal efficiency, or also with intertemporal distribution are essentially questions of the second best. The general question of the instruments by which government can alter the intertemporal allocation of intertemporal resources is a question to which we shall return.

1.3 Distinctions between Static and Intertemporal Resource Allocations

There are important similarities, then, between the problems of intertemporal allocation of natural resources and the problems of the static allocation of resources. But there are important differences, and although it is important to keep in mind the similarities, the differences may be even more important.

1.3.1 Over-saving

The earliest noticed potentially important difference between the efficiency with which markets allocate resources in a static environment and in a dynamic context arises from the possibility of over-saving, and was first noticed by Samuelson [1958]. He pointed out that if individuals saved for their retirement, and had to hold their savings in the form of capital, then the interest rate could be lower than the rate of growth, and thus the economy would be dynamically inefficient.[1] However, although this possibility has been extensively discussed in the literature (see Cass and Yaari [1966], Diamond[1965], it is probably of little practical importance, for at least two reasons: first, so long as there is money (or an alternative, non-produced store of wealth), this particular type of over-saving cannot occur. Second, few economies seem to have experienced interest rates for a sustained period anywhere near as low as that required for over-saving to be considered an important possibility.

In the recent literature on exhaustible resources, it has been shown that in a competitive economy, over-saving of capital cannot occur, but the possibility of excessive conservationism, i.e., using up natural resources too slowly, cannot be ruled out. (See Stiglitz [1974a], Dasgupta [1973].)

1.3.2 Absence of Futures Markets

This arises from the second, and probably more important difference between a temporal and intertemporal allocation of resources, the absence of futures markets, extending infinitely far into the future. For without these markets, it is impossible to be sure within a competitive (market)

---

[1] That is, some generation could be made better off without making any other generation worse off. Consider the special case of zero growth rate. Then, clearly, if the marginal return on investment is negative, one could reduce the capital stock today and increase output in all subsequent periods, and still be better off. The generalization of this principle to growing economies is discussed by Solow [1962].

economy that the price for the natural resource or capital goods will be correct; there is a large variety of initial prices all of which are consistent with market equilibrium (the return to all assets being the same) in the short run. In some classes of models, including the simple natural resource models investigated so far (Stiglitz [1974b]), only one of these leads to balanced growth; (see also Shell-Stiglitz [1967]; the earliest discussion of this point is due to Hahn [1966], but in other models there may be an infinite number of initial prices, all of which converge to balanced growth (Shell, Sidrauski, and Stiglitz [1967]).

The question naturally arises of whether the problems associated with the absence of futures markets are any more serious for natural resources than for other long lived assets. Although there has been some discussion of this question in the literature (Stiglitz [1974b], Nordhaus [1974]) there appears to be no convincing answer. We discuss three aspects of this problem below.

The first concerns the observed relative instability of resource markets. Many commodity markets are characterized by large fluctuations in prices. These fluctuations may have many explanations. They could, for instance, represent the fluctuations which would occur as a result of new information about future demands and supplies. Or the instability could reflect an instability in the short run dynamics of the economy, the period by period adjustment of the market, not the long run dynamics of which we have been speaking so far. (In this case, it is not the absence of futures markets which is the central feature; on the other hand, to the extent that the absence of futures markets forces individuals to rely on expectations about futures prices, which will in general depend on current prices, if these expectations are elastic with respect to current prices there will be some tendency for instability in the markets. Which of these factors, or indeed, whether some other factor, is primarily responsible for these fluctuations in prices is a moot question.[1]

A second aspect relates to the irreversible nature of the consumption of a natural resource. Mistakes in the kind or amount of capital goods constructed can, at a cost, be corrected. But once a natural resource is used up, it cannot be "disused." This irreversibility makes it imperative that the resources be used correctly, and, it is argued, provides an important argument for conservationism. As it stands, at least, this argument is spurious; the argument might be correct if we were to use our entire stock of the resource up, and then discover that it was necessary for production, i.e., no substitute for it existed and it was essential in some important production process. But in general, the views about the future needs will be continuously revised, and as the stock dwindles, and its price rises it will become apparent that the stock has no substitute; this will lead to a higher price of the stock, and a lower level of consumption. The stock will be conserved for the future. The cost of the lack of foresight is economic waste, using the resource inefficiently (i.e., relative to how it would have been used with perfect

---

[1]See Stiglitz [1974] and Nordhaus [1974].

foresight), but there is no reason to believe that this cost is of a different order of magnitude than the costs associated with the inefficiencies involved in building the wrong amounts or kinds of capital goods.

A third aspect of the problem which has received some discussion arises from the fact that conservationism entails a reduction in consumption by the present generation, and an increase in consumption in later generations. Since the individuals who benefit from conservationism are not alive now, there is no way the market mechanism can assure that their interests are protected, so it is argued. And since the lifetime of the resources exceeds that of any single individual, no individual within his lifetime has the incentive to make sure that the allocation is correct. These arguments are also spurious. Investors need not cut down the trees to realize a return from owning a forest, to give but one example: all they need to do is sell the trees to someone who realizes that eventually they can be cut down. Individuals who assist in the intertemporal allocation of natural resources, by projecting demands and supplies, obtain their return from the capital gains they make on these investments, the difference between the buying price (today) and the selling price (tomorrow)[1] in just the same way that an entrepreneur obtains his return from buying in one market, where the commodity is cheap, and selling in another, where it is more expensive. And in this argument, the fact that the individuals who buy from the speculator at some date in the future are not alive at the time the speculator makes his purchase is irrelevant.

---

[1] And a symmetric argument holds for selling short.

In short, the absence of futures markets means that the analysis of the optimality of the market allocation of resources cannot be simply applied to the analysis of the intertemporal allocation of natural resources. It does not, of course, mean that there is any systematic bias in the market allocation of resources; to determine this we need to say a great deal more about the structure of these markets, a question to which we shall turn in later sections. It does, however, raise an important welfare-theoretic question in the evaluation of the efficiency of the market allocation. There is sometimes a tendency to compare the allocation which would have prevalied if there were a complete set of markets with that which prevails under an incomplete set of markets, and, if the two differ, to conclude that our actual market allocation is inefficient. Although there is a sense in which such a statement is true, it does not provide us much insight with respect to policy. There may be good reasons that a complete set of markets does not exist, e.g. the functioning of such a set of markets might be prohibitively costly. Thus, the relevant question from a policy view, is an evaluation of (a) whether the market attains a constrained pareto optimality, i.e., is efficient, subject to the constraint that only the given markets function and (b) whether there are too few (or too many) markets in operations. For instance, there is some argument for indicative planning based on the absence of futures markets. (See Meade [1970]). These issues will become somewhat clearer in the context of the analysis of the absence of risk markets, below.

1.3.3 Absence of Risk Markets

A closely related set of differences between the conventional static welfare analysis and the analysis of resource markets arises from the absence of a complete set of risk markets.

There are a number of important sources of uncertainty in natural resource markets. First, there is uncertainty about the total stock of the resource (say oil). Secondly, there is technological uncertainty concerning the costs of extracting the oil (this is closely related to the first, in that the amount of recoverable oil is a function of the technology of extracting oil). Thirdly, there is uncertainty concerning the demand for oil. This is a reflection of uncertainty concerning the price of presently available substitutes, and uncertainty concerning the development of new substitutes. These are all considerations which, to a competitor in, say, the oil industry, become reflected in the price at which he can sell oil at some date in the future. Note that these are all real, social risks. The opportunitycost of using more oil today is using less oil in the future; the point is that we do not know what this opportunity cost really is. No matter at what rate we consume oil, there is some contingency in which, ex post, the opportunity cost to using oil today turns out to be very high, and there is some contingency in which it turns out to be too low. The crucial problem is in the balancing of these various possibilities.

In addition to these social risks, there are private risks: an individual who purchases a particular piece of land may not know whether there is oil underneath it, or the costs of extracting it. But the discovery or non-discovery of oil at that particular location may have no (or at most a negligible) effect on society's view of the total available supply of the resource. These are private risks, in the sense that, in principle, an individual could diversify out of these risks; that is, in a well functioning securities market, since the "event" "discovery of oil at this particular location" is uncorrelated with the occurrence of any other event, the ownership shares in this risk would be sufficiently diversified that all individuals would be risk neutral with respect to the outcome. In fact, most of the risks which were discussed in the earlier literature, arguing that the oil industry was risky, focused on these private risks.

To be sure that one obtains a pareto optimal allocation of resources requires a complete set of risk markets, as we commented above. In the absence of a complete set of risk markets, two problems arise: there will be exchange inefficiency, i.e., the marginal rates of substitution of consumption in two different states may differ for two different individuals, and the market may not provide the correct signals (or incentives) for investing in different risky assets. A judgment of how badly the market does, whether there are systematic distortions, and how one ought to correct those distortions depends on one's view of the capital market, a subject about which there is little agreement.

There is one widespread view that for private risks of the kind described above, the market is efficient, that given the diversification possibilities within the market, the market does act as if it were risk neutral with respect to such risks.

The evidence on this is far from unambiguous, and there are those who believe that even for such risks, the market acts in a risk averse manner. One reason which has been put forward for explaining this is that because of imperfect information and the costs associated with becoming informed, individuals tend to own a relatively small number of securities, i.e., they are incompletely diversified (see Stiglitz [1975b]).

As in the discussion of the consequences of the absence of a complete set of futures markets, one must be careful to make the appropriate welfare comparisons. Even more than there, it is inappropriate to compare an allocation with a complete set of insurance markets with the allocation which actually results. So long as individual enterpreneurs still must bear the risks, and they cannot diversify out of them, then this risk bearing is a real social cost. The appropriate policy question in that context is, are there institutional modifications which will allow a better diversification of risk.

There is also considerable controversy on the question of whether the risk markets, even if efficient in diversifying risks, provide the correct signals for investing. (In this particular context, not-consuming oil is equivalent to investing.) It was argued in the earlier literature (e.g., Diamond [1967]) that, even if there were not a complete set of futures markets, if firms acted competitively, then, under certain admitted restrictive technological conditions (that the pattern of outputs across the states of nature remained unchanged as the investment level changed -- multiplicative uncertainty -- and firms had no choices other than the level of investment) the economy would attain a constrained pareto optimum i.e., the decentralized market would attain the same results that a central planner who was constrained to distributing output as a linear function of the outputs of the different firms would obtain. It was subsequently shown (Stiglitz [1974c], Hart [1974]) that an essential assumption in that analysis was that there was only a single commodity, so relative prices between commodities did not depend on the investment decisions. If there is more than one commodity -- as there must be in any model analysing natural resources -- then the competitive market allocation will not in general be pareto optimal.

Whether there is any systematic bias in the market allocation is a more difficult question to answer. For a rather special model, where the mean-variance model is applicable (implying either a joint normal distribution of returns or quadratic utility functions), Stiglitz [1972] and Jensen and Long [1974] have established that the market systematically is biased against risky investments, provided that firms attempt to maximize their stock market value. Whether this is the correct objective function of the firm has been a subject of some extensive debate (Stiglitz [1972a], Leland [1974]), but the earlier results of Stiglitz, that there is some presumption that firms will underinvest in risky assets, still remains valid provided the typical owner is, over his lifetime, planning to be a net seller of shares.

There have not been any systematic attempts to apply these general results to natural resources. There is a widespread belief that because the price of oil is high when there is oil scarcity (a bad state of nature from the point of view of society) and the price of oil is low when there is an overabundance of oil (a good state of nature from the point of view of society) returns to speculators in oil are negatively correlated with social welfare in general; the consequence of this is to reinforce the impression that the market will not take the correct actions with respect to investing in oil (not consuming oil). But the absence of a systematic analysis of this is one of the deficiencies we hope to remedy in the next stage of our work.

2. The Structure of Efficient Allocations of an Exhaustible Resource

In this section we shall abstract from the distributional issues raised earlier and concentrate on the behavior of market prices in an efficient, but not necessarily optimal intertemporal allocation. To simplify matters, we shall postulate a single extractive resource industry, assume no uncertainty and perfect capital markets.

A necessary condition for intertemporal efficiency is equilibrium in asset markets. That is, the sum of rentals plus capital gains, less depreciation, must be equal for all assets, so that all assets yield the same rate of return, which we call the market interest rate. This is frequently termed the arbitrage condition.

If there were a complete set of forward markets, competition would equalize rates of return on all assets, and the arbitrage condition would hold at every point in time. We shall imagine instead that at each date, t, there are only forward markets for the next period, which we will denote by $(t + \theta)$. This insures that all prices are known at $t + \theta$, and all rates of return may be determined.

This sequence of markets might be called the myopic economy. Flow and asset markets clear, so that resources are efficiently allocated in the present and current investment is divided among competing activities to maximize return in the next period.

We shall contrast the allocation determined by a myopic economy with the efficient allocation that would obtain if futures markets were complete, or the resource were socially managed.

It is shown that in the absence of uncertainty, a myopic economy can be directed to achieve an efficient intertemporal allocation by merely specifying the correct initial price of the resource. For simplicity, we compare the myopic and efficient allocations for the case of a resource with zero extraction cost.[1]

The Myopic Economy

Denote by p_t the competitive spot price of the resource stock at time t. Consider an individual who has p_t dollars at time t. Given the market interest rate, r_t, he may buy, say, bonds and be assured of $(1 + r_t\theta)p_t$ dollars at time $(t + \theta)$. Since the resource stock does not depreciate or earn rentals during the period $(t, t + \theta)$, the return to holding the resource is determined completely by its change in price. The forward market for the resource stock will clear at the price $p_{t+\theta}$ for which

$$(1) \qquad p_{t+\theta} = (1 + r_t\theta)p_t \quad .$$

Taking the limit at $\theta \to 0$, we obtain the movement of the spot price of an exhaustible resource as[2]

$$(2) \qquad \frac{\dot{p}_t}{p_t} = r_t \quad .$$

---

[1]We use extraction cost as a generic term which includes all factor costs required to produce the resource.

[2]A dot over a variable signifies time differentiation.

It would not be an exaggeration to regard equation (2) as the fundamental principle of exhaustible resources. The only way that a given stock of such a resource can yield a return to its owner is by appreciating in value. It follows that under competitive conditions it is the rate of capital gains enjoyed by the resource that must equal the return earned in holding any other asset. The question also draws attention to the following important feature. Even if r_t were constant over time, so long as it is not zero the spot price of the exhaustible resource cannot remain constant over time. Consequently, so long as an economy possesses exhaustible resources, as well as resources that earn rentals, one cannot contemplate the existence of steady states with constant factor prices. Spot prices cannot remain constant over time under any circumstance.

We have derived the arbitrage equation (2) as a condition describing stock equilibrium in the market for assets. It will prove instructive to derive it as well as a condition of flow equilibrium in the market for the exhaustible resource. For simplicity let us assume zero extraction costs. Let p_t^R denote the spot price of a flow of the resource at t and let p_t continue to denote the price of a unit of stock at t. Assume that there are competitive markets for both the flows of the resource and its stocks. Suppose an individual owns S_t units of the stock at t. Then the spot value of his asset at t is p_tS_t. Under competitive conditions this must be equal to the maximum present value, calculated at t, that he can earn by extracting the resource over time. If it were greater, no one would wish to extract. If it were less, no one would wish to hold on to the stock at all. Let $R_\tau (\tau \geq t)$ denote an extraction policy. For simplicity suppose that the return to holding the numeraire asset is constant, $(r > 0)$. Given the price paths p_t and p_t^R the owner will wish to select that time profile of extraction R_τ which will maximize his present value of profits. It follows then that

$$(3) \qquad p_t S_t = \max_{(R_\tau)} \int_t^\infty p_\tau^R R_\tau e^{-r(\tau-t)} d\tau \quad ,$$

where

$$(4) \qquad \int_t^\infty R_\tau d\tau = S_t$$

Let $R_\tau^*(\tau \geq t)$ be the solution to the maximization problem (3). Then differentiating both sides of equation (3) with respect to t one obtains

$$(5) \qquad \dot{p}_t S_t + p_t \dot{S}_t = r p_t S_t - p_t^R R_t^* \ .$$

From (4) we have that $\dot{S}_t = -R_t$. Using this in equation (5) yields

$$(6) \qquad (\dot{p}_t - r p_t) = (p_t - p_t^R) \frac{R_t^*}{S_t}$$

But under competitive conditions the stock price must equal the flow price. Thus $p_t = p_t^R$. It follows from equation (6) that $\dot{p}_t / p_t = r$; which is precisely the arbitrage condition. But it is also clear that with the price of the resource rising like compound interest at the rate r the value of the stock (and, therefore, the maximum present value of sales) is independent of the actual extraction policy; as long, of course, as the entire stock is exhausted over the future. In other words R_t^* is not uniquely given. To see this, given that $\dot{p}_t/p_t = r$, we have on integration that $p_\tau = p_t e^{r(\tau-t)} = p_t^R e^{r(\tau-t)}$. Using this in the RHS of equation (3) reduces the equation to

$$(7) \qquad p_t S_t = p_t \int_t^\infty R_\tau d\tau \quad .$$

All this is merely a round-about method of saying that if the price of the resource rises at the compound rate r, owners of resource stocks will be indifferent at the margin between extracting (and selling the resource flow) and holding at each instant. It is then possible to imagine the overall rate of extraction of the resource just equal to the competitive demand at the current price, with the result that the market for the resource flow clears at each instant. We shall confirm this presently. But before undertaking to do this there are two observations that need to be made. First, an inspection of equation (2) suggests that even if the entire time profile of r_t is known the arbitrage condition merely dictates the percentage rate at which the spot price of the resource must change; it does not provide any instruction for determining the price level at which to commence. In other words, we are in need of an initial condition. Second, even though we have assumed extraction costs to be nil, p_t is not nil. That is to say, if the resource is of use its competitive price will be positive (and rising of the percentage rate r_t) even though "production" costs are nil. Stating it another way, with extraction costs assumed away, the entire value of a stock of the resource is composed of the flow of services it can provide. This is why the competitive value of a pool of oil of a deposit of coal is often referred to in the literature as its royalty value. It is worth re-emphasizing this point, and in particular the fact that under competitive conditions the spot price of the exhaustible resource rises at the compound rate r_t. This tilt in the competitive price path is an inescapable feature of an exhaustible resource with negligible extraction costs and, to look at the matter from another point of view, a necessary condition for an efficient utilization of the resource. In other words, the fact that the spot price is rising over time is in itself no evidence of a growing monopoly power of the owners of the resource.

"(OPEC supporters)...seem to be rejecting the play of free market forces in determining prices. In such a market, the price of a product is closely related to the cost of producing the last unit of supply that is demanded by a buyer. No one anywhere in the world is pumping oil that costs \$10 a barrel to 'produce.' The cost of bringing up a barrel ranges from 10¢ in Saudi Arabia to 60¢ in Venezuela to \$3 or so in the U.S. OPEC's defenders seem to have the notion that somehow market forces have never properly recognized the value of oil, that its price always should have been higher. This tosses rational economic analysis out of the window."[1]

---

[1]Time magazine, October 14, 1974, page 36.

As we have already remarked, exhaustible resources present several special features that are often overlooked. The foregoing problem is an apt example of the kind of argument one must avoid.

Let us now introduce the demand side of the picture. Suppose that the market demand curve for the flow of the resource (R_t) at t is given by the function $D(p_t,t)$. If the resource is a factor of production like ores and fossil fuels, then D is a derived demand curve. To conduct the analysis in the simplest possible manner suppose that the demand curve does not shift over time and, in particular, that

$$(8) \qquad R_t = D(p_t,t) = p_t^{-\frac{1}{\alpha}} ,$$

where

$$\alpha > 0 .$$

In other words, demand is iso-elastic. Let r be constant (> 0). Then integrating equation (II.2) yields

$$(9) \qquad p_t = p_0 e^{rt}$$

Notice that (9) describes the movement of the spot price of the resource. Since the return on holding the numeraire asset is by assumption r, the present value price of the exhaustible resource is $p_t e^{-rt} = p_0$. It is constant over time.

Let us commence at $t = 0$, when the total stock of the resource is S_0. Competitive conditions prevail and the resource flow market is assumed to clear at each date. Using (9) in equation (8) then yields

$$(10) \qquad R_t = p_0^{-\frac{1}{\alpha}} e^{-\frac{r}{\alpha}t} .$$

The rate of utilization of the resource falls at a constant percentage rate r/α. This is so because as the market price rises the current rate of utilization falls along the demand curve. Now recall that we are aiming here to describe a sequence of momentary equilibria, where at each instant the market for the resouce flow equilibrates and where the asset market equilibrates as well. If we were to postulate the existence of a complete set of competitive futures markets, then the equilibrium price system would certainly satisfy equations (8) - (10). But the arguments leading to equations (8) - (10) have not depended on the existence of a complete set of forward markets. We have supposed merely that at each date both the asset market and the market for the flow of the resource equilibrates. This observation alone should warn us that such a competitive process may well misbehave. We can confirm this.

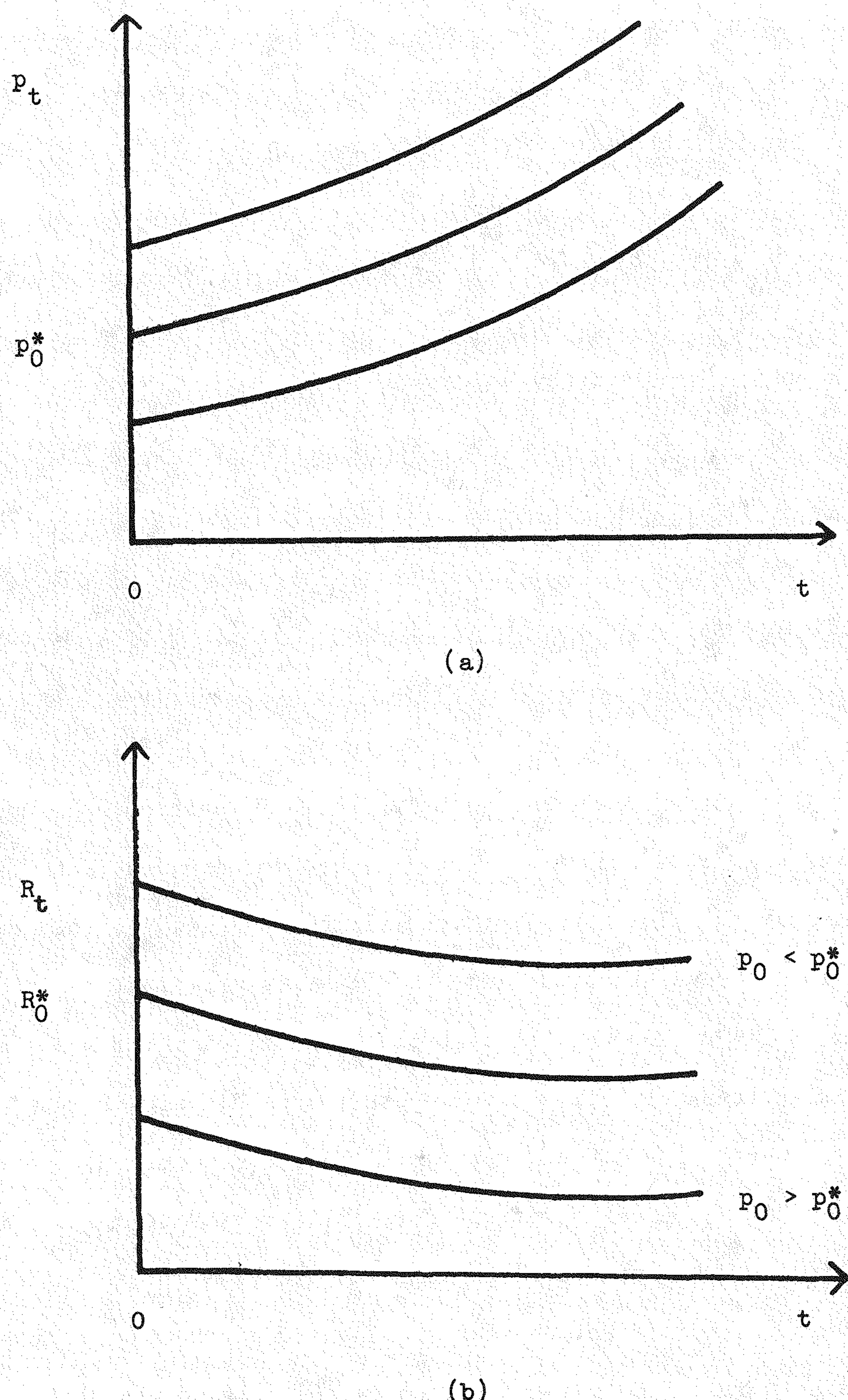
In formal terms one should notice that none of the arguments establishing the conditions for momentary equilibrium allows us to determine the initial price p_0. It is as yet an unknown of the problem. And it will remain so unless we impose further structure on the construction: namely, the assumption of a complete set of forward markets. But given that in the world as we know it, such a complete set of markets does not exist it is essential to see the implications of this indeterminacy of p_0. Toward this let us integrate (10) to yield

$$\int_0^\infty R_t dt = \int_0^\infty p_0^{-\frac{1}{\alpha}} e^{-\frac{r}{\alpha}t} dt = \frac{\alpha}{r} p_0^{-\frac{1}{\alpha}} \tag{11}$$

The total stock initially is by assumption S_0. Define p_0^* as

$$p_0^* = \left(\frac{rS_0}{\alpha}\right)^{-\alpha} \tag{12}$$

It is then immediate from equation (11) that if $p_0 = p_0^*$ the competitive process will just exhaust the resource in the very long run. The rate of extraction will be $R_t = (rS_0/\alpha)e^{(-r/\alpha)t}$ (see diagrams II.1(a) and (b)). But suppose the initial price is "wrongly" set. In particular, suppose that $p_0 > p_0^*$. In this case the price will be "too high" at each date and consequently the flow of extraction will be "too low" at each date; "too low," in the sense that the integral of sales will fall short of the total stock, S_0. Such an outcome is patently inefficient, since a marginally lower initial price would allow the process to yield a higher rate of extraction at each instant. But what would be the economic motivation for certain owners of resource to refrain forever from extracting? It is this. Since any given stock appreciates in value at the rate r a typical resource owner is indifferent between storing and extracting. Furthermore, if there is no terminal date for the economy there is no end to the process. Therefore a typical resource owner may as well never extract. It follows



<u>Figure 1</u> Price and extraction paths for different initial prices.

that if $p_0 > p_0^*$, in order for the sequence of momentary equilibria to be sustained, there will always be some part of the stock that never gets exhausted.[1] Notice that this feature would not occur if resource owners were far-sighted and in particular, if a complete set of futures markets were to exist. It is arising because we are postulating a competitive process in which agents are myopic. Resources owners expect the price of the resource to rise at the rate r at each instant, and their expectations are indeed fulfilled at each instant. This is why the arbitrage equation is often termed the myopic rule.

What if $p_0 < p_0^*$? The situation here is different. The spot price is "too low," and the rate of extraction is "too high." Clearly then, were the process to persist the resource would be exhausted in finite time. The question then is whether the process could persist. Probably not. If resource owners see the exhaustion of the resource in sight (i.e., they realize that the resource will be exhausted in finite time), there will be tremendous gains to be had from hoarding the resource until the date of exhaustion and selling after this date. But as soon as traders attempt to buy up stocks the spot price of the resource would jump in value in order to restore the asset market equilibrium. In other words such a disequilibrium behavior as we have sketched should get the economy off any path that has $p_0 < p_0^*$. This is, of course, not a conclusive argument; merely the articulation of a belief that traders act on the basis not only of price signals but other signals as well. In particular one would imagine that if traders recognized that there is positive demand for

---

[1] The possibility of a systematic underutilization of resources along a competitive process with no terminal date was originally noted in Malinvand [1953], Koopmans [1957] and Samuelson [1958]. In a different context the point was raised also by Stahn [1966].

the resource flow even at very high prices, that the current price is not very high and that the stock is very nearly depleted, they would use these pieces of information in making decision. But as we have remarked, the reverse case (i.e., when $p_0 > p_0^*$) could easily persist. Resource owners would recognize that there will be no "shortages" under the process in the near future. But unless they are very far sighted they would not know that a certain portion of the original stock will never get extracted.

Although extremely simple, the foregoing example has brought out an important form of market failure. In the absence of a complete set of futures markets an indefinite sequence of momentary equilibria could rather readily lead to an inefficient outcome, one where the outcome results in too much conservation, not too little. The result would not be surprising. In order to calculate the correct price today one needs to know where the economy ought to move to in the long run. In the absence of an announced target it should not be surprising that there are errors all along the way. One way for the entire set of future possibilities to be taken into account at the initial date is for there to be a complete set of forward markets. Another is to contemplate a planning board, announcing national prices (or shadow prices) to guide the pace of extraction. Let us look into this now.

A Socially Managed Exhaustible Resource

Let $R = D(p)$ be the market demand curve for the resource flow. Since by assumption $D'(p) < 0$ we can invert the function. Thus

$p = D^{-1}(R)$. Let us write this as $p = B(R)$. The gross rate of consumers surplus at R is by definition $\int_0^R B(R')dR'$. The return on the numeraire asset is by assumption $r\ (>0)$. We shall imagine that the planning board is concerned with the pace of extraction of the resource (of size S_0 at $t = 0$) with a view to maximizing the present discounted value of the flow of gross consumers' surplus.[1] We are, then, supposing for vividness that the resource is socially managed. The planning problem then is this:

Choose an extraction path $R_t (t \geq 0)$ with a view to maximizing

$$(13) \qquad \int_0^\infty [e^{-rt} \int_0^{R_t} B(R')dR']dt$$

subject to the constraints $S_t = S_0 - \int_0^t R_\tau d\tau$ and $R_t, S_t \geq 0$ for $t \geq 0$. Write $u(R) = \int_0^R B(R')dR'$. We can now express the objective function as

$$(14) \qquad \int_0^\infty e^{-rt} u(R_t)dt \quad .$$

Since $B(R) > 0$ for $R \geq 0$ we have $u'(R) > 0$ for $R \geq 0$. Under this interpretation, it is trivial to define a Pareto efficient extraction policy. First, a definition of an <u>in</u>efficient extraction policy. For simplicity suppose that we are concerned solely with the class of all <u>continuous</u> extraction paths. By a <u>feasible</u> extraction policy R_t we shall

---

[1]The assumptions implicit in this objective are discussed elsewhere in this report. Among other things one is supposing that r is judged by the planners to be the appropriate rate of discount.

mean one which satisfies the constraints in (13). It is then natural to say that a feasible policy R'_t is <u>inefficient</u> if there is a feasible policy R_t such that $R_t \geq R'_t$ for all $t \geq 0$ and such that for some finite interval of time $R_t > R'_t$. It is equally natural to regard a policy R_t as <u>efficient</u> if it is <u>not</u> inefficient. Given one definition it follows immediately that an extraction policy R_t (≥ 0) is efficient if and only if $S_0 = \int_0^\infty R_t dt$.[1] But problem (II.13) is not concerned with only efficient extraction policies, but rather with determining the best policy. Given that $u'(R) > 0$ we know at once that the optimum extraction policy (if one exists) will in fact be efficient.[2]

Locating the condition which the optimum extraction policy must satisfy is a simple enough matter. It is the obvious analogue of condition (2). Let us simplify matters and suppose that along the optimal policy the rate of extraction is always positive. Then the condition we shall get is the analogue of (9), which states that along the optimal path the marginal social valuation of resource utilization at each date is constant when looked at from date $t = 0$. In other words one must have

(15) $$e^{-rt}u'(R_t) = \lambda \quad (> 0)$$

But by hypothesis, if R is the rate of extraction, $u'(R)$ is the market clearing price for the resource flow. If the planners call forth for a rate of extraction R_t, the flow will be sold in the market at the price $u'(R_t)$. In other words, the present value price of the resource must

---

[1]We stress that at this point we are contemplating a world with no uncertainty.

[2]In what follows we shall not bother with the question of the existence of a solution.

remain constant along the optimum extraction path. We are now home. For writing $p_t = u'(R_t)$ and differentiating equation (15) with respect to time yields the condition $\dot{p}_t/p_t = r$ which is all very familiar. The remaining bits of the analysis do not bear repeating. Since the optimum policy must be efficient it must satisfy both equation (15) and the condition $\int_0^\infty R_t dt = S_0$. There is no indeterminacy in the initial price level (or equivalently, the initial rate of extraction). This brings us back to our earlier observation. A complete set of futures markets, or, alternatively, the conscious efforts of a planning board could in principle be relied upon to coordinate flows and stocks of an exhaustible resource in such a manner so as to result in an efficient outcome. But in the absence of either, a competitive process, even if it were to sustain an equilibrium at each instant of time, cannot be relied upon to achieve this result. Perhaps we should not have expected anything else.

Extraction Costs

It is time to introduce extraction costs into the picture. Any reasonable account of the economics of exhaustible resources must allow for the fact that resources are utilized in the process of extraction. There are several features to be considered. The first is particularly relevant for resources such as oil and natural gas, where the recoverable stock depends on the rates of extraction. If the rates of extraction from a given field are unduly high, a certain portion of the stock gets dissipated and cannot be recovered. This would appear to have occurred in several of the oil fields in the eastern states of the U.S. until recently. It is plain enough that this aspect of extraction costs is difficult to formalize in a simple manner. A loose approach would be to suppose that the average cost of extraction is an increasing function of the rate of extraction. Second, extraction technology would appear to improve with time. Third, the average cost of extraction from a given deposit would seem to depend on the stock remaining. In particular, one would suppose that the marginal cost of extraction would increase as the stock diminishes, given that one is, as it were, digging deeper into the ground with less of the stock left.

We shall analyze the influence of extraction costs in what is otherwise the model presented in the previous section. For simplicity of exposition assume that there is a single pool of the resource. Let C denote the cost of extracting the resource at the rate R when the stock size is S. A simple form of such a cost function would then be

(16) $$C = f(t)R_t g(R_t)G(S_t)$$

where

$$f'(t) \leq 0 \quad , \quad g'(R_t) > 0$$

and

$$G'(S) \leq 0 \quad .$$

Such a cost function needs hardly any explanation. It patently captures the three features that we have encountered.

It is simplest to regard extraction costs as "transport" costs involved in moving the resource from the source to the market. This alone suggests that the effect of extraction costs is to drive a wedge between the price of the unextracted resource and the price of the extracted resource, and in particular, that the latter price is higher. It is then immediate that the differences between these two prices is the marginal cost of extraction. Let p_t continue to denote the competitive spot price of the unextracted resource and let q_t be the competitive spot price of the extracted resource. Then

$$(17) \qquad q_t = p_t + \frac{\partial C}{\partial R_t}$$

and it is q_t that determines the volume of the resource flow that clears the market.[1] Consider now the arbitrage equation. As before, let r be the market rate of interest. Under competitive conditions the return on holding the marginal unit of the stock consists of two components. The first consists of the capital gains that the stock enjoys. This in fact constituted the entire return in the model of the previous section. The second consists in the <u>reduction</u> in future extraction costs due to the fact that this marginal unit has been stored, and not extracted. The arbitrage condition is a statement regarding the equality of these two rates of return at each instant. That is,

$$(18) \qquad \frac{\dot{p}_t}{p_t} - \frac{\frac{\partial C}{\partial S_t}}{p_t} = r \quad .$$

---

[1]In the previous section we supposed that $C = 0$ and hence $\partial C/\partial R = 0$. This implied that $q_t = p_t$.

From (18) it is now clear that $\dot{p}_t/p_t < r$, and also clear why it must be so. There are gains to be had in storing over and above capital appreciation.[1] Furthermore, equation (17) suggests that if the marginal cost of extraction declines rapidly over time, because of innovations in the technology of extraction, it is possible for q_t to decline over time for awhile. But it does bring out clearly that the price of the extracted resource consists of two components: the marginal cost of extraction and the royalty price. Thus, while it is true that "...the price of (an exhaustible resource) is closely related to the cost of producing the last unit of supply that is demanded by the buyer,"[2] this "close" relationship is in fact an extremely complicated one and not much can be asserted without solving these foregoing equations. In the absence of extraction costs obtaining the explicit price trajectory was an easy enough matter. Not so when extraction costs functions assume complicated forms. Imagine then that the average cost of extraction is a constant, $b(>0)$. In terms of (16) this is tantamount to assuming that $f(t)g(R_t)G(S_t) = b$. From equation (18) it is then immediate that

(19) $$\frac{\dot{p}_t}{p_t} = r \quad ,$$

and from equation (II.17) that

(20) $$q_t = p_t + b \quad .$$

---

[1]One should, perhaps, point out that by "storage" we mean "not extracting." So long as there are costs of extraction it never pays under competitive conditions to extract and to store above ground. This is because $r > 0$.

[2]cf. footnote 3.

From (19) we have $p_t = p_0 e^{rt}$ and consequently

$$(21) \qquad q_t = b + p_0 e^{rt} \quad .$$

We need to estimate the "correct" initial royalty price p_0^*. Continuing to assume an iso-elastic demand curve it follows that [1]

$$(22) \qquad q_t = b + p_0 e^{rt} = R_t^{-\alpha}$$

and, therefore,

$$(23) \qquad R_t = (b + p_0 e^{rt})^{-\frac{1}{\alpha}}$$

For this extraction path to exhaust the resource completely it is required that

$$(24) \qquad S_0 = \int_0^\infty R_t \, dt = \int_0^\infty (b + p_0 e^{rt})^{-\frac{1}{\alpha}} dt \quad .$$

The "correct" initial price p_0^* is the solution of equation (24). Let us, by way of illustration, assume $\alpha = 1$, (the elasticity of demand is unity). Then integrating (24) yields

$$(25) \qquad S_0 = \frac{1}{br} \log \frac{b + p_0^*}{p_0^*}$$

[1] For convenience of exposition we shall continue to assume that there is a positive demand at any price, no matter how high the price. The iso-elastic demand curve will meet this requirement. Thus in fact $R_t > 0$ for all t along the competitive path.

therefore,

(26) $$p_0^* = \frac{b}{(e^{brS_0} - 1)}$$

Using (20) and (26) one obtains

(27) $$q_t = b + \frac{be^{rt}}{(e^{brS_0} - 1)} \quad .$$

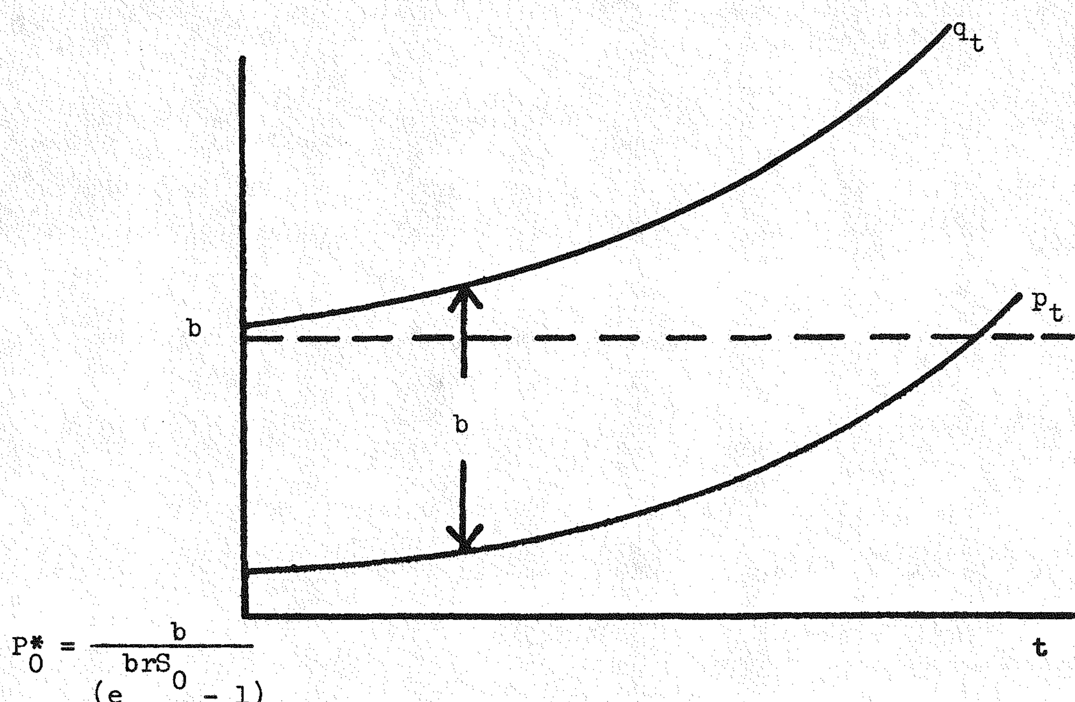


Figure 2

Equation (27) exposes the two components of the competitive price of the extracted resource in an extremely tidy manner. Analytically, the interesting situation is one where S_0 is "large." In this case p_0^* is "small" and consequently $q_t \doteq b$ initially. The price of the unextracted resource (p_t) rises exponentially at the rate r, but the price of the extracted resource (q_t) does not. In fact q_t rises at a variable rate less than r, for $\dot{q}_t/q_t = rp_0^* e^{rt}/(b + p_0^* e^{rt}) < r$. In other words, q_t does grow, but given that by assumption p_0^* is "small," the royalty component is negligible during the early years, and the extraction cost component dominates. This makes good intuitive sense. If the stock is large the fact that the resource is exhaustible is not of much concern. It is much like a conventional commodity whose unit cost of production is b. But with time the royalty component begins to dominate, since $q_t \doteq (be^{rt})/(e^{brS_0} - 1)$ for large t. In other words, with time the fact that the resource is exhaustible begins to bite, and the production cost becomes a negligible part of its price. Indeed, in the long run $q_t \doteq p_t$ and the spot price of the extracted resource grows roughly at the rate r. (See diagram 2) In short, it all depends on how large the initial stock is. But how "large" is large? To get a feel for this note that since $(p_0^*/b) = (1/(e^{brS_0} - 1)$, (equation (26)), if, say, $brS_0 > 6$ then $(p_0^*/b) < 0.01$; that is, the initial price of the unextracted resource ought to be less than 1% of the marginal cost of extraction. Suppose, by way of illustration that $S_0 = 250$ (i.e., the total stock is 250 times the current rate of consumption of the resource), that

$r = 0.05$ per annum, and that by normalization, $b = 1$. In this case $brS_0 = 12.5$. Using equations (21), (26) and (27) it can be computed that it will be about 130 years until brS_t falls to a value as low as b.[1] In other words, the equilibrium price, q_t, will roughly equal b until the remaining stock is less than forty times the current annual rate of extraction.

Now it must be admitted that unchanging demand conditions (as caught in (8)) are unlikely. One would imagine that, if anything, demand will grow. This will work against the orders of magnitude we have just presented, in the sense that for a given initial stock the gap between q_t and b will be larger if demand for the resource is assumed to increase, rather than stay constant. This brings us back to our earlier observation that the relationship between q_t and the marginal cost of extraction depends on a number of complicated features. Roughly speaking q_t hovers near this cost of extraction so long as the stock is large in some sense. Not so otherwise.

Let us complicate matters somewhat and suppose that there are different deposits of the same resource; the difference being in unit extraction costs. In particular suppose there are two deposits with unit extraction costs b_1 and b_2, $(b_2 > b_1)$.[2] Let p_{1t} and p_{2t} be the spot prices of the unextracted resource at these two deposits, and let q_{1t} and q_{2t} denote their extracted spot prices. The first point to

---

[1] These figures are taken from Kay and Mirrlees [1975].

[2] One might wish to suppose that the second deposit is less accessible or some geological difference that makes extraction cheaper at the first deposit.

note now is that under competitive conditions the two deposits will not be mined simultaneously over any interval of time. Suppose they were. Then owners of the two deposits will be indifferent between storing and extracting over this interval. Consequently,

$$(28) \qquad \frac{\dot{p}_{1t}}{p_{1t}} = \frac{\dot{p}_{2t}}{p_{2t}} = r \quad .$$

But since the two deposits contain identical products, $q_{1t} = q_{2t}$. This implies that

$$(29) \qquad b_1 + p_{1t} = b_2 + p_{2t} \quad .$$

Now equations (28) and (29) are patently inconsistent with one another. Therefore the two deposits will not be mined simultaneously but instead will be mined sequentially. The precise sequence is exactly as one would imagine; the better quality deposit will be mined first until it is exhausted, and the lower quality deposit will be mined subsequently. This is precisely what considerations of efficiency dictate. Given that $r > 0$ it makes sense to delay mining the higher cost deposit. Formally, the argument runs as follows. So long as stocks of both deposits are positive, equation (28) will hold in order that the asset market clears. For an initial period (T) the second deposit will be found unprofitable to mine. That is

$$(30) \qquad q_t = b_1 + p_{1t} < b_2 + p_{2t} \quad \text{for} \quad 0 \leq t < T \quad .$$

During this period owners of the first deposit undercut the price of the second one. The owners of the second deposit store, and do not find it profitable to extract, since the market price of the extracted resource does not cover the higher cost of extraction (b_2). Meanwhile q_t grows continuously (this follows from (28) and (30)), and the extraction rate falls.[1] At T the first deposit is exhausted and there is a switch to the second deposit. The point to note now is that q_t must be continuous at T. For suppose that at T the extracted price were to jump to a higher value. Then since we are supposing that resource owners would know of this jump at T, owners of the first deposit could increase their profits by ceasing production just before T and producing after T. But by hypothesis they cease operating after T. Consequently there cannot be a discontinuous increase in the price at T. Nor can there be a discontinuous fall at T since otherwise owners of the second deposit would find it profitable to enter production before T. This establishes the continuity of q_t at T. From date T the second deposit takes over and it is as though there is a single deposit. The story from then on is the same as the one we have already constructed for the competitive extraction from a single quality deposit. In particular for $t \geq T$ one will have $q_t = q_{2t} = b_2 + p_{2t}$. The continuity of the market price for the extracted resource (q_t) at T is an extremely valuable piece of information. Diagram 3 portrays the price trajectories. An important point to bear in mind is that during (0,T), while the prices

---

[1]This last follows directly from our assumption of an unchanging downward sloping demand curve.

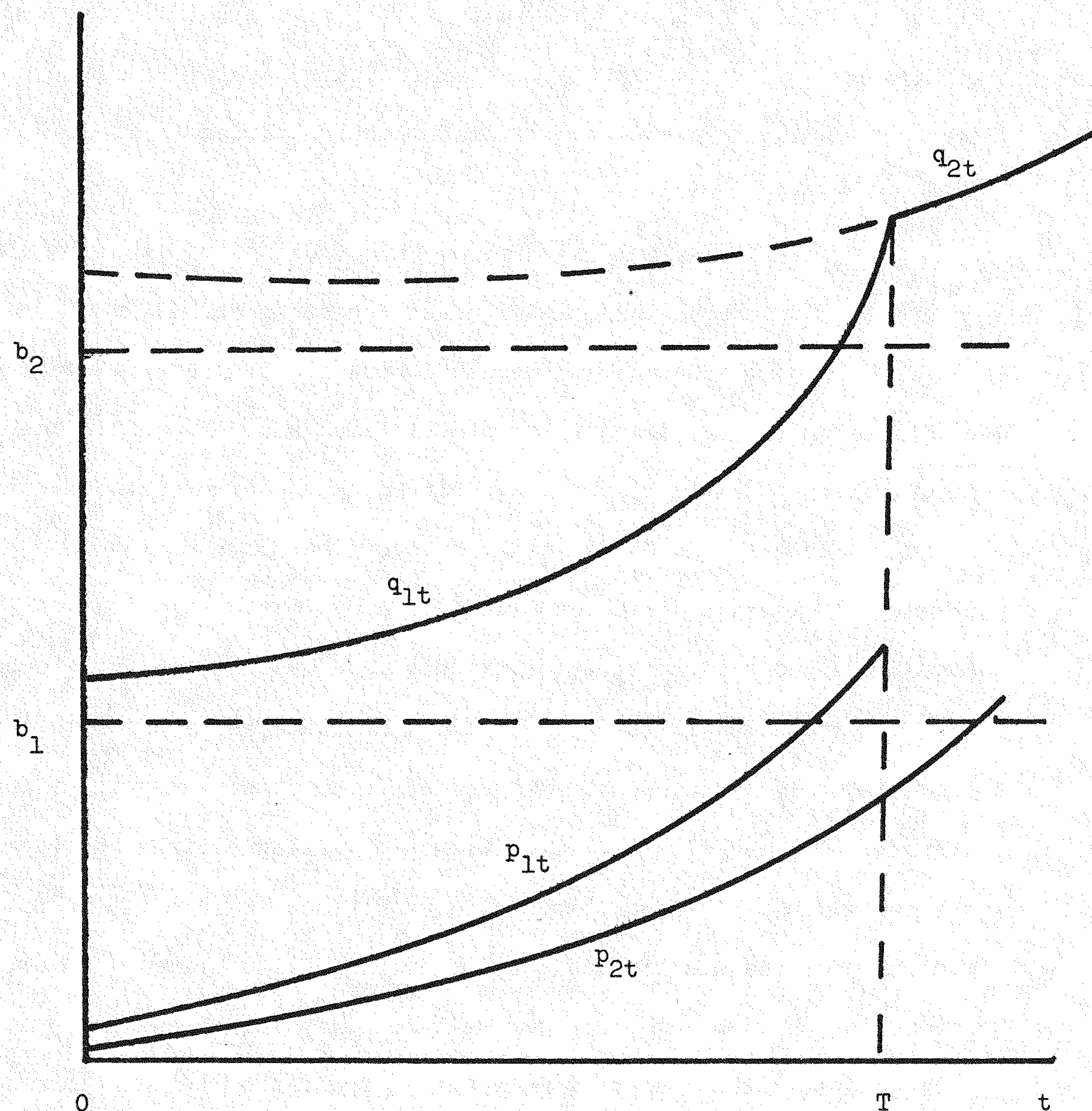


Figure 3

of a unit stock in both deposits rise at the same rate (equation (28)), the price <u>levels</u> are by no means the same. Even common sense suggests that the unit price of the second deposit will be lower during $(0,T)$. It is less valuable given that it is costlier to mine. Since demand conditions are, by assumption, unchanging, the rate of extraction under competitive conditions is a declining function of time. Figure portrays this.

There is a limiting case of this model which is of considerable interest. Suppose that the lower grade deposit contains in effect an unlimited stock. Controlled nuclear fusion or solar energy would provide a motivation for our being interested in such a model. Until they come about, an approximation might be provided by the production of energy by breeder reactors. In other words, suppose that in addition to a given deposit of an exhaustible resource (with unit extraction cost at b_1) there is a known technology for producting a substitute product whose unit cost of production is b_2.[1] Since the substitute is by assumption a conventionally produced good its market price under competitive conditions is precisely b_2.[2] Our earlier discussion will have made it clear what the price movement under competitive conditions will be. Let p_t denote the price of the unextracted resource and q_t its price after extraction. Given that $b_1 < b_2$ the deposit will be mixed for an initial period $(0,T)$ while the substitute product will not be found profitable to manufacture.

---

[1] In effect we are supposing that $S_0 = \infty$. Nordhaus [1974] has christened it as the "backstop technology."

[2] In order to check that this fits in with our discussion so far, note from equation (27) that $q_t \to b$ if $S_0 \to \infty$.

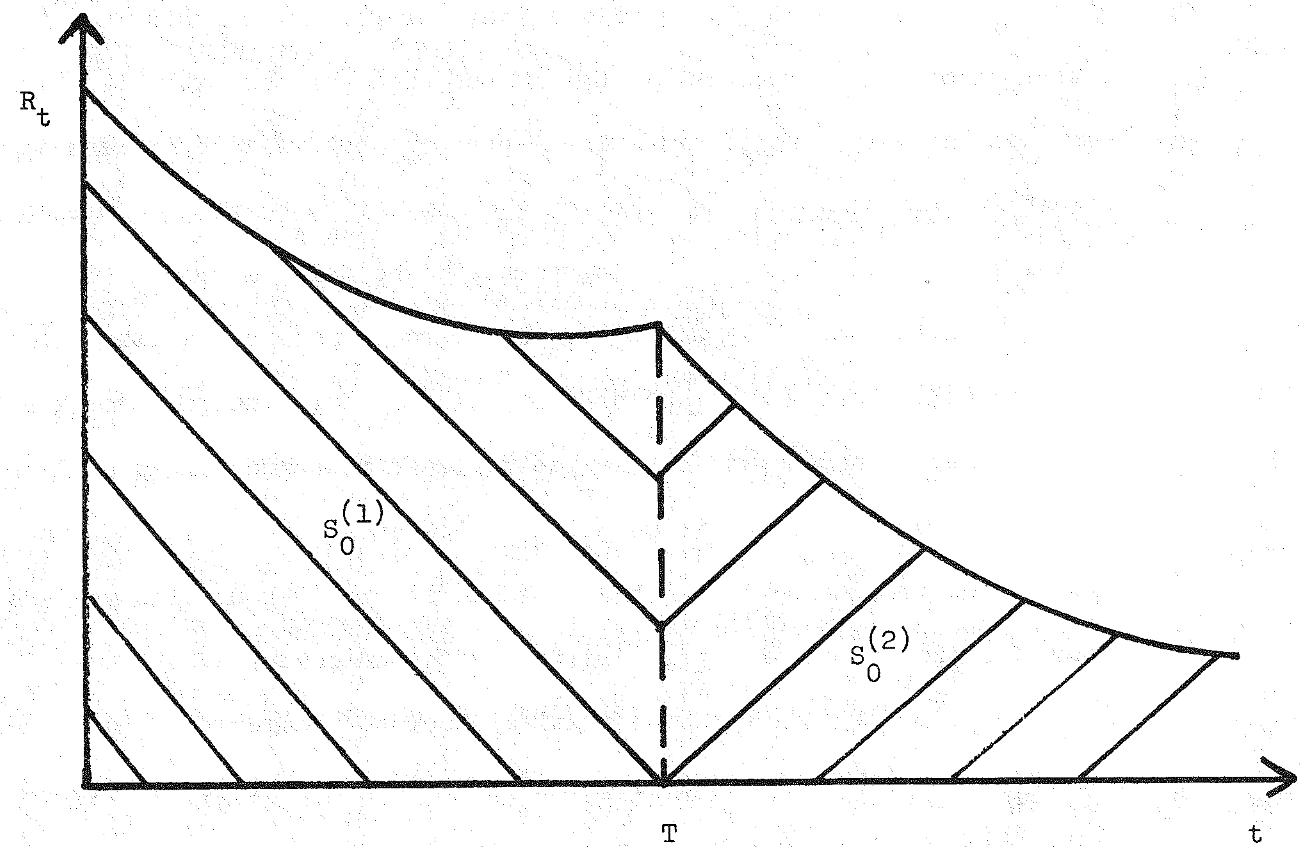


Figure 4

Figures 5 and 6 depict the time profiles of the competitive price and the rate of production of the resource and its substitute. (T is the date at which the resource is exhausted and the backstop technology is introduced.)

While exceedingly intuitive, the foregoing result is important to bear in mind. It reminds us that if the unit cost of extraction is less than the unit cost of producing a product that is more or less a perfect substitute, then the substitute ought not to be manufactured initially.[1] The backstop technology ought to be held in reserve until the resource is completely depleted. Of course, under competitive conditions the backstop technology <u>will</u> be held in reserve until the resource runs out, because so long as stocks are not depleted resource owners will be able to undercut the competitive price of the substitute product. The result is of importance for another reason: it provides us with well-defined bounds for the competitive price of the extracted resource, for $b_1 < q_t \leq b_2$. The larger is the initial stock the longer will q_t hover near b_1 (see (27)) and consequently the longer will it be before the backstop technology makes its appearance.

Let us, by way of illustration, see what orders of magnitude may be involved. Suppose demand is unchanging and iso-elastic and that the price elasticity is unity. Suppose that $S_0^{(1)} = 50$ (i.e., total

---

[1]We are now making use of the efficiency properties of the intertemporal competitive equilibrium, hence the term "ought" in the sentence. Furthermore, we are assuming the production of substitutes can be described by purely convex cost functions.

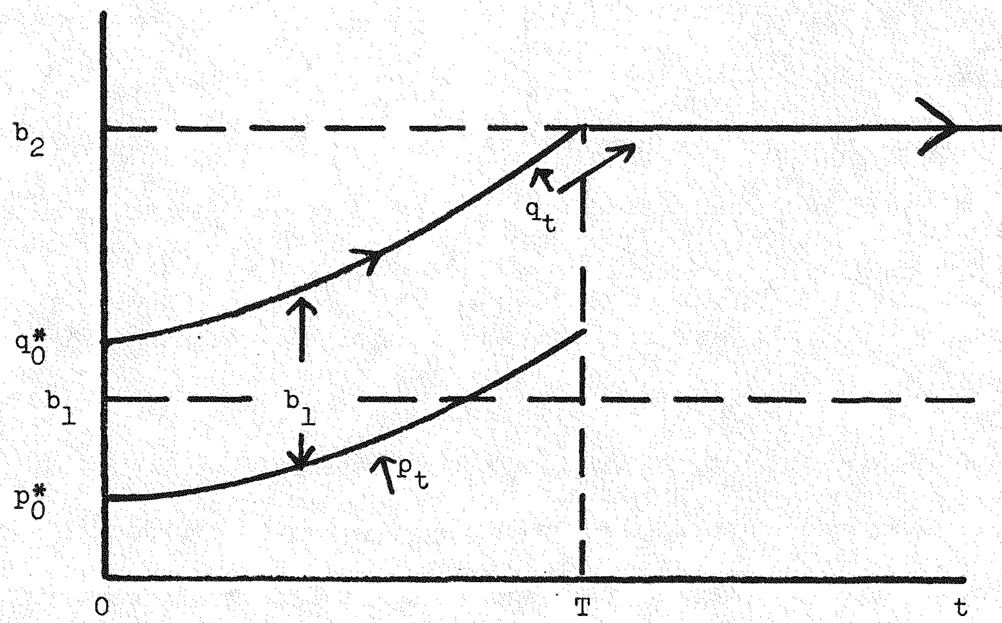


Figure 5

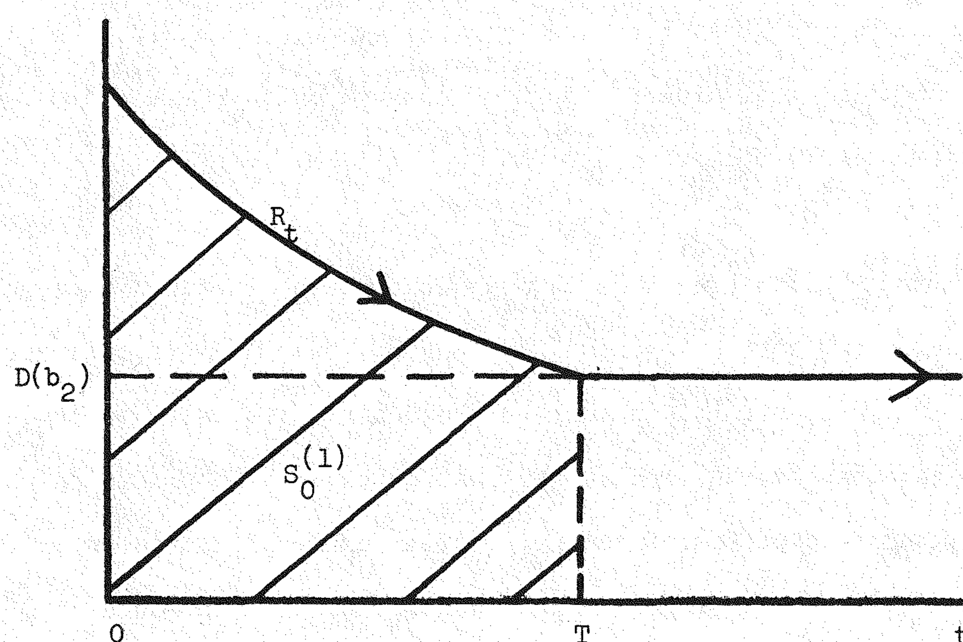


Figure 6

stock is 50 times current rate of extraction) and $r = 0.05$ per annum. By normalization set $b_1 = 1$ and suppose $b_2 = 10$.[1] It is then simple to confirm that $p_0^* < 0.01$ and that $T \simeq 180$ years. Of course, these conclusions cannot be taken seriously. If nothing else, demand is expected to grow, and the numerical results will not be as striking then. But they do suggest that the royalty component of the price of many exhaustible resources may well be rather small in initial years.

It is occasionally claimed that the competitive price of an exhaustible resource is the cost of producing a substitute product. In a widely read popular report on the recent four-fold price increase of Arab oil this view was attributed to OPEC defenders:

> "Nearly every OPEC member..., rejects the notion that the price of oil is now too high. 'What do they mean by high?' asks Iran's Minister of the Interior... incredulously. He reasons that the price is about equal to what it would cost to obtain an alternative form of energy, such as gas produced from coal. Thus he... insist(s) that $9.70 per barrel is a fair price."[2]

In the long run, when the resource is near exhaustion, the argument would seem reasonable, but not until then. So long as there is a considerable gap between b_1 and b_2, and so long as the stock is large, the competitive price of an exhaustible resource is well below the cost of producing the substitute. Any attempt to set the price of an exhaustible resource roughly equal to the cost of producing a substitute product would result in excessive conservation.

---

[1] This is not an entirely fictitious figure. $2.00 a barrel is a rough average figure for extracting crude oil and it is occasionally said that the cost of producing shale oil is the equivalent of $20.00 per barrel of crude. Oil shale, while not exactly unlimited in quantity is pretty much like a backstop technology.

[2] _Time_ Magazine, October 14, 1974, page 36.

Conclusion

Let us sum up:

(1) The competitive (efficiency) price of an exhaustible resource is the sum of the marginal cost of extraction and the price of the unextracted resource (the royalty price). In particular, the extracted price exceeds the marginal cost of extraction and is not equal to it.

(2) If the cost of extraction of a resource is independent of the stock, the royalty price will, under competitive conditions, grow at a compound rate. This rate of increase will equal the rate of return on holding the numeraire good.

(3) An implication of (2) is that the present value royalty price of a resource is constant over time.

(4) An implication of (1) and (2) is that the price of the extracted resource will grow at a lower rate (and may fall initially) than the rate of increase in the royalty price. But in the long run the two prices will grow at the same rate.

(5) If the marginal cost of extraction is a declining function of the stock, the percentage rate of change in the competitive royalty price of a resource is less than the rate of return holding the numerair, the difference being the reduction in future extraction costs as a result of storing the marginal unit of the stock.

(6) A sequence of momentary equilibria (one along which at each date both the asset market and the market for the extracted resource is in competitive equilibrium) cannot be relied upon to ensure an efficient

utilization of an exhaustible resource. In particular, if the initial royalty price is set "too high" there will be excessive conservation. A fraction of the initial stock will never get extracted. But the 'market' may never discover its erroneous ways.

(7) If there is a complete set of futures markets, or if there is a planning board looking far into the future, such a market failure in principle can be circumvented.

(8) An implication of (1) is that if the stock is large an extracted resource is much like a conventional produced commodity; in that its price roughly equals the marginal cost of extraction (i.e., the royalty price is negligible). As the stock diminishes, if there are no substitutes in sight the fact that the resource is exhaustible becomes important. At this stage the price of an extracted resource roughly equals the royalty price, which is far in excess of the cost of extraction.

(9) The existence of a substitute product for an exhaustible resource does not imply that the competitive (efficiency) price of the extracted resource is the cost of producing the substitute. Indeed if the cost of extraction is less than the cost of producing the substitute product the competitive price of the extracted resource lies between these cost figures. If the stock is large the price will be roughly equal to the cost of extraction (i.e., the royalty price is negligible). It is only when the resource is near exhaustion that the price rises to the cost of producing the substitute.

(10) An implication of (9) is that the fact that a substitute has been discovered does not provide an argument for producing it initially. So

long as the production cost of the substitute exceeds the extraction cost of the exhaustible resource the substitute ought not to make its appearance, and under fully competitive conditions it will not make its appearance initially. The resource and the substitute will not be utilized simultaneously, but sequentially.

3. Sources of Market Distortions in the Intertemporal Allocation of Exhaustible Resources

3.1 Introduction

We are used to thinking of market prices as giving the appropriate signals for whether an industry should expand or contract, what inputs to use in its productive processes, etc. However, the list of necessary and sufficient conditions for market efficiency is quite extensive, and it does not require a great deal of investigation to identify areas in which existing markets for natural resources are imperfect.

There is probably no sector of the economy which is not affected to some extent by imperfect markets. To show that market imperfections provide cause for selective government intervention in natural resource markets, one must be able to establish that, for some reason, the market imperfections are more important, or have more important consequences, in the energy industry; and that there are specific government instruments which can be used to alleviate these problems.

We have already noted that efficiency is only a precursor for optimality, and government intervention in efficient markets may be justified for distributional objectives. In the following discussion, we shall presume that the private rate of interest equals the social consumption rate of interest, so that the efficient rate of resource utilization is socially optimal as well. If private and social rates of time preference were different, this would require that the government exercise the proper mix of monetary and fiscal measures to equalize the market rate of interest and the social consumption rate of discount. Conditions under which government actions can achieve this objective are discussed in Arrow and Kurz [1970].

We shall also assume that those policies invoked to achieve distributional objectives preserve efficiency; and we shall abstract from the existence of any other policies which may result in dead weight loss.

The presumption of equality between social and private rates of interest and the absence of distortionary taxes and regulations allows us to regard market efficiency as a first-best policy objective. Having noted the conditions for market efficiency in an idealized economy, we may now direct our attention to determine the adequacy of existing markets for natural resources.

The following areas have been identified as potentially important sources of deviations between social and private returns. We shall describe the existing literature pertaining to these sources of market imperfections and indicate, where possible, the direction of the market bias relative to the efficient allocation.

3.2 The Value of Information and Exploration of the Resource Stock

The uncertainties and risks which were a focus of discussion in section I give rise to attempts to acquire information to reduce the uncertanties. The traditional welfare analysis of the competitive economy assumes that the state of information is invariant, or at least that it is not affected by the actions of the individuals. The recent studies of Spence [1973], Rothschild-Stiglitz [1975], Stiglitz [1975c], Grossman [1975], and others have established the possibility of non-pareto optimal equilibria in the presence of imperfect information. Several of these studies have also questioned the appropriateness of the perfectly competitive assumptions in the presence of imperfect information.

There have been a few attempts, in the works of Peterson [1975], Gilbert [1975], and Stiglitz [1975d] to apply the general principles to exploration for new resources. The approach taken in this section is first to discuss in general terms the ways in which competitive markets may fail to properly allocate investment in the generation of information. We shall then examine the uncertain characteristics of the endowment of an exhaustible natural resource and consider the social value of information about these characteristics. Thirdly, we shall analyze specific instances of market failure in the production of information about the resource stock.

In this discussion, we concentrate on exploration information concerning the physical characteristics of the resource. Analysis of uncertainty and the value of information about future demand, import supply, and technological developments will not be discussed at this time.

3.2.1 Information and Economic Analysis

The fundamental problem in the economic theory of information arises from the public good character of information; the use of information by one individual does not deny its use by others. In fact, like national defense and electromagnetic waves, information is one of the few pure public goods. Samuelson [1954] has shown that the market price of a pure public good will underestimate its social value and a competitive market will underinvest in its production. Indeed, this is the classical argument for government subsidization of basic research.

Hirshleifer [1971] has argued that the market may, in some instances, overinvest in research. As an example, consider the stock market. The social value of information about a particular security is the increase in consumer surplus resulting from a change in the pattern of investment brought about by the information. If the information relates to a new technology for producing a substitute energy source, the social value of the information may be very large. On the other hand, the information may be limited to next quarter's earnings of a firm, and will have little bearing on future production and consumption decisions, in which case its social value is negligible. The private value of the information is the speculative gain which may be realized by purchasing securities whose prices will rise when the information is released, and selling short over-priced securities. It is clearly possible for the private gain to exceed the social value of the information, and under such conditions, the market will provide excessive incentives for the production of information.

Stiglitz [1975e] has reconciled these conflicting views by calling attention to an implicit assumption in the Hirshleifer hypothesis; namely, that an informed individual exercises monopoly power in the market for information. Indeed, if two or more informed individuals compete for speculative profits, they will bid up under-priced securities and bid down over-priced securities until profits are reduced to zero.

Green [1973] and Grossman [1975] have carried the argument further by demonstrating that there need be no direct interaction between agents, since prices are a natural source of externalities. A movement in prices may signal uninformed individuals that the economy has received some input of information, and in some situations, the prices may be a perfect surrogate for the information.

When there is a competitive market for information, and prices adjust without friction, only the owners of an asset may benefit from information bearing on the value of the asset, since price adjustments will eliminate speculative gains. The analogy of the incentives to produce information with incentives to produce public goods may be misleading. If prices adjust freely, and an entrepreneur cannot internalize the benefits of his research, then the market offers no incentive for research, not just too little incentive. Conversely, the incentives for research may be excessive if imperfect market conditions permit speculative gains.

Information may also take on the aspect of a public inferior good, and a competitive economy may overinvest in the production of this kind of information. For example, information may serve to discriminate one group from another, and the market-determined investment in discrimination may exceed the optimal level.

We have identified three problems in the economics of exploration information:

(1) As a public good, there is insufficient private incentive for its production.

(2) As a means to appropriate rents by being better informed, there may be too much private incentive for its production.

(3) As a pure screening device, a means to distinguish a good or service, there may be too much incentive for its production.

The nature of the market bias in the generation of information depends on the specific environment in which the information is produced. We will assess the market bias in the generation of information concerning the endowment of a natural resource, and offer reasons to suspect that in some activities the market allocation of investment in the generation of information may be excessive, while in others the market may underinvest in the production of information.

3.2.2 The Social Value of Information about an Exhaustible Resource

There are at least two areas in which information may increase the social value of an exhaustible resource.

(1) The total stock of the resource.

This is important in determining the appropriate rate at which the resource should be consumed. If we think there is less oil than there is in fact, then we will be too conservation minded; if we think there is more oil than there is in fact, we might consume it too quickly, leaving our descendants with an inadequate stock of oil. To be more precise, information bearing on the total stock of the resource as a function of extraction cost has social value in determining the appropriate utilization rate.

(2) The cost of extraction from a particular tract of land.

This is broadly defined to include the location of resource deposits, since absence of a deposit may be viewed as existence of a deposit with an infinite extraction cost. Conceivably, on the basis of statistical information, we could obtain a good estimate of the total stock of oil, without knowing where it is located, or the cost of extraction from any location. For instance, it might turn out that 50% of a certain type of geological structure contains oil, but until we explore a particular geological structure, we do not know whether that particular geological structure contains oil. If the cost of extraction at each location were known with certainty, the optimal pattern of extraction would entail extraction from deposits in order of increasing cost. This is obvious, since by extracting the cheaper oil first, one can postpone the higher costs of extraction, and the present discounted value of total extraction costs will thus be lowered. This aspect of exploration can be viewed as screening different parcels of land, distinguishing those with lower expected extraction costs from those with higher expected costs.

In the analysis of efficient resource allocation with perfect information, the rate of change of price consistent with momentary equilibrium depended on extraction costs, and the price level necessary for intertemporal efficiency depended on the total resource stock. Given uncertainty, these equilibrium conditions cannot be assured. Failure to maintain the extraction path consistent with the true values of these parameters of the resource stock does not imply inefficiency, since resources must be expended to identify their actual values.

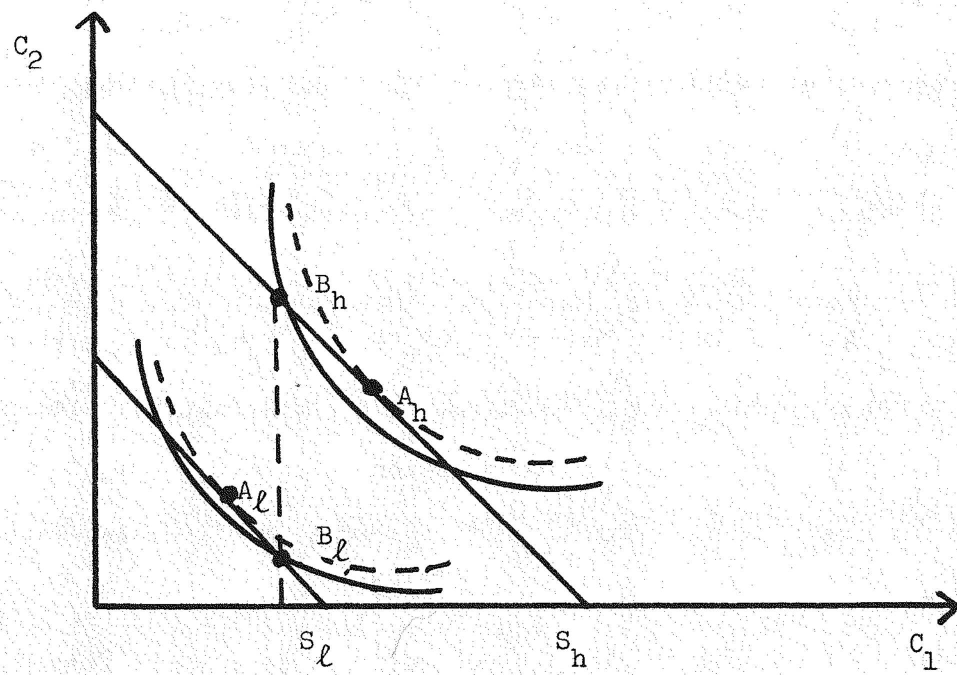
Consider the value of information on the total resource stock. Imagine that the U.S. economy is constrained to follow a policy of zero imports, and the only source of uncertainty is, say, the amount of oil which may be off the Atlantic coast. For simplicity, we suppose that there are only two time periods of interest -- now and then -- and the total U.S. oil stock may be either high if the Atlantic OCS is productive, or low if it is not.

The potential of the Atlantic OCS can be "proved," but this is a costly process. The trade-off between the cost and value of this information is illustrated in figures 1 (a) and (b). Consumption in the first period is plotted on the horizontal axes and consumption in the second period is plotted on the vertical axes. The straight lines indicate that the resource endowment, either high, S_h, or low, S_ℓ, may be allocated between the two periods in any proportion. The curved lines represent points of indifference between the consumption pairs, (C_1, C_2). Indifference curves farther from the origin represent higher levels of social welfare.

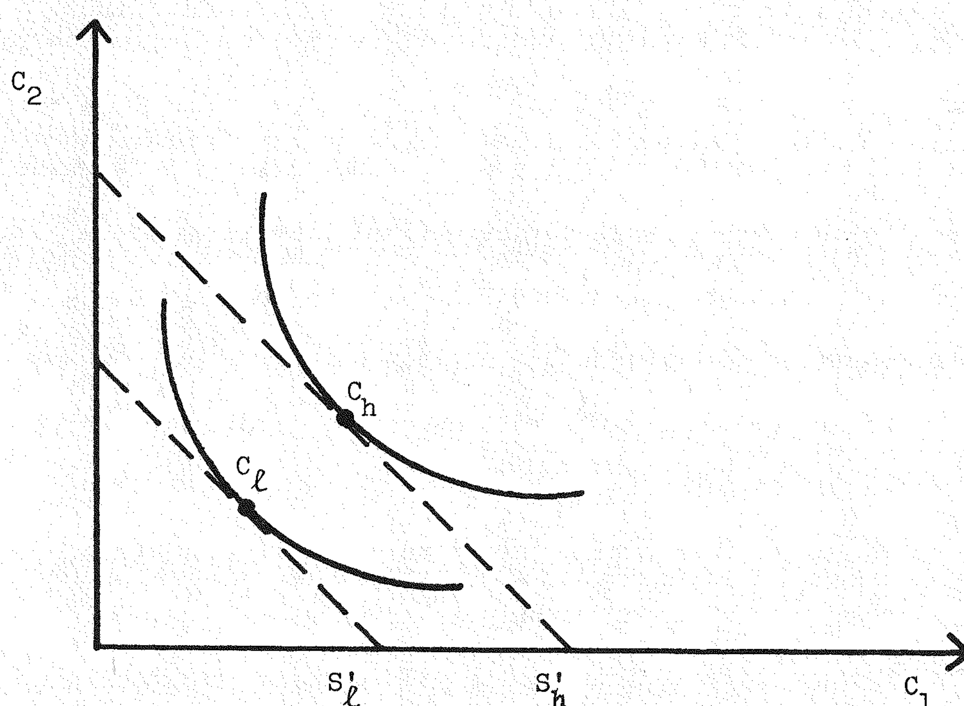
In the Utopian world of perfect information at zero cost, consumption could be allocated to maximize welfare contingent on the stock, e.g., points A_ℓ and A_h in Figure 1 (a). In the absence of perfect information, first-period consumption cannot be made contingent on the true value of the stock. The optimal plan calls for choice of a consumption level in the first period such that expected social welfare is maximized. In general, the chosen C_1 will not be optimal for either value of the stock. This is illustrated in Figure 1 (a); the points B_ℓ and B_h are the resulting consumption pairs.

There are two components to the cost of information about the amount of Atlantic OCS resources:

- The cost of advancing exploratory activity that could be postponed if information were unnecessary.
- The possible efficiency loss if increase in the pace of exploration brings diminishing returns.



(a)



(b)

Figure 1

The first cost component clearly depends on the opportunity cost of capital, and if the discount rate were zero, this cost would be zero.

The cost of exploration can be considered as a tax on the total available stock of the resource. That is, exploration reveals the true size of the resource, but consumes some in the process. This is illustrated in Figure 1 (b), where $S'_\ell < S_\ell$ and $S'_h < S_h$. The points C_ℓ and C_h correspond to the allocations contingent on S'_ℓ and S'_h respectively.

Whether the social value of exploration of the Atlantic OCS exceeds the social cost depends on such factors as the consumption elasticity of substitution between the different time periods, the variance of estimates in the total resource stock, and the cost of proving the potential of the Atlantic OCS.

3.2.3 Market Failure in the Production of Exploration Information

Information which relates only to the total size of the resource base is clearly a public good, since the information affects the choices of all consumers and producers of the resource. Hence the private value of exploration may underestimate its social benefits. Indeed, if the firm is a price-taker and marginal extraction costs are constant, the value of information to the firm is precisely zero. The firm will allocate production solely as a function of price changes and demand. If the exploration information generated by the firm does not affect prices, it will not influence production decisions and hence has no value.

Exploration may be motivated by potential cost reductions from locating lower cost deposits and postponing extraction from higher cost deposits. Cost reduction is a socially productive activity and there is no evidence of faulty market incentives for this component of exploration. Market incentives may distort the timing and location of exploration when there are spill-overs in exploration information, and when the sale of mineral rights provides opportunities to appropriate profits.

Spill-overs in information (both discovery and non-discovery of oil at a particular location convey information about the probability of discovery or non-discovery at other locations) are likely to result in too little post bidding exploration, unless there is unitized exploration. This is a consequence of the fact that, unless an entire geological structure is owned by a single producer (or there is unitization of exploration and drilling) success or failure at one point conveys information about the presence of oil at other points. This is an illustration of the public good nature of information. The information has a social value, to the extent that it affects drilling decisions of others, but these social returns cannot easily be appropriated by the individual drilling the well. Similarly, the individual has an incentive to postpone his exploration, (and consequently extraction) in the hope that someone else will do some exploratory drilling, the information from which he can "steal."

A simple example illustrates the nature of the problem. We assume there are two contiguous sites, owned by different individuals, A and B. Each assumes that drilling has a positive expected return.

However, A is best off, in an expected utility sense, if B drills first, and then A does; conversely for B. If B does not drill, A may be better off drilling or not drilling depending on the cost of waiting and the change in the costs of drilling. If the cost of waiting is not very great, A may be better off waiting if B does not drill, and B may be better off waiting if A does not drill. Thus, there is an "equilibrium" with no exploratory drilling the first period. Nonetheless, exploratory drilling on the part of one firm would be socially profitable.

The second source of distortion arises from expenditures to appropriate rents, perhaps most evident in the closed-bid government leasing program for outercontinental shelf oil tracts. The government leases its off-shore oil tracts with a reservation price, but the reservation price is sufficiently low that the expected value of the oil exceeds the reservation price. Thus, if there were a single bidder, the government would be giving away some rents. But to obtain these rents, the bidder must know what to bid, and this requires that he form an estimate of the amount of oil. More generally the expected return of the uninformed bidders, in an optimal bidding strategy, is zero, but for the informed bidders it is a function of the number of bidders. If there is a cost of becoming informed (of going from the "uninformed state" to the "informed state") this will determine the equilibrium number of bidders. But the expenditure on this exploratory activity is determined not by the social returns, but simply by the desire to obtain these rents. Thus, if the government puts up the tracts for lease at a date earlier than the socially optimal date for development, these tracts will be explored too early; since, clearly, optimality requires that this type of exploratory activity be postponed until immediately before exploitation of the tract.

Moreover, since the (expected) returns are a function of the (expected) number of bidders, there may be an incentive to explore and bid on relatively out-of-the-way tracts.

In our earlier discussion, we noted that there was a social return from screening different tracts of oil by the cost of extracting the oil. Assume all oil tracts had exactly the same amount of oil, and differed <u>only</u> by the cost of extraction. Then, whether it is optimal to screen depends simply on the cost of screening relative to the present discounted value of the savings from postponing postponable drilling expenses.

If there is a market for tracts of oil, then whether or not it is socially desirable to screen according to quality, it may pay for individuals both to find out about the quality of their wells (the cost of extraction) and to make this information available. Thus, the presence of the asset market today for oil wells, the oil from which will be extracted at some date in the future, leads both to the acquisition of this information when it is socially inefficient, and even when it is socially desirable, to its acquisition earlier than is socially desirable. (Such might be the case, for instance, if it were known that the well had a sufficiently high extraction cost, that it should not be used for at least five years, but the magnitude in excess of that amount was not known. Social optimality requires that if this well is to be screened for extraction costs, it not be screened for five years, i.e., the earliest date at which it might be desirable to extract the oil.)

A second source of distortion in the timing -- and from a practical point of view probably the more important source -- arises from the so-called diligence requirements on government leases: owners of the lease are required to develop the lease within a given period, or else they lose the lease. Thus, the condition for developing the lease within that period is only that the price exceed the cost of extraction which is obviously not socially optimal. It also means that it does not pay firms to screen for the cost of extraction, since they will have to develop the well anyway within a short period of time.

Note that a resource monopolist would value information on the size of the total resource stock as well as the extraction costs of deposits. In the next section, we discuss how a monopolist fails to efficiently allocate the resource. However, it is interesting that a monopolist might better manage exploration of the resource; although the monopolist's exploration pattern is not necessarily efficient because the monopolist's objectives are not consistent with maximization of social welfare.

3.3 Imperfect Competition

We have seen that efficient utilization of exhaustible resources requires that:

(a) the rate of return on resource deposits equals the rate of return on alternative assets, and

(b) the resource stock is exhausted when its price equals the price of a substitute source of supply.

This suggests that we examine the effects of imperfect competition in two areas: the market for the exhaustible resource and the market for substitute sources of supply.

The qualitative work of Hotelling [1931], Sweeney [1975], and Stiglitz [1975f] provides a structure for quantitative comparisons of the intertemporal extraction of resources under competitive and monopoly ownership of the resource stock. They focus on a comparison of profit-maximizing behavior under competitive and monopoly ownership. The effects of changes in the market for substitute sources is not explicitly considered.

We shall summarize the results obtained under the assumption that the marginal extraction cost of the resource is constant. The more complicated case of variable extraction costs is discussed in Sweeney [1975].

The competitor seeks to maximize profits taking prices as given:

$$\pi = \int_0^\infty \{P(t)q(t) - Cq(t)\}e^{-rt}dt$$

subject to

(1) $$\int_0^\infty q(t)dt \leq S_0$$

where

π = profits

$P(t)$ = price at time t

$q(t)$ = production at time t

C = unit extraction costs (assumed constant)

r = market interest rate

S_0 = resource stock owned by the firm .

The competitor will choose the production path, $q(t)$, which maximizes profits. Assume there are many competitors who face the same extraction cost, C. The marginal profit of production at time t' (in present value terms) is

(2) $$\{P(t') - C\}e^{-rt'}$$

for each competitor.

If the marginal profit of production at t' exceeds the marginal profit at any other time, all competitors will choose to produce their entire stocks at time t'. This flooding of the market will depress prices and lower marginal profits. Assuming sufficient price elasticity and complete futures markets, prices will adjust so that

(3) $$(P(t') - C)e^{-rt'} = \text{constant}$$

for all t'. Differentiating (3.2), we have

(4) $$\frac{\dot{P}(t)}{P(t) - C} = r$$

Equation (4) states that the net rent on resource deposits must rise at the market rate of interest. Prices in an efficient competitive economy will adjust so that the net rate of return to holding resource deposits is the same as the net rate of return on other capital stocks. With complete futures markets or perfect expectations, the absolute resource price level will adjust to allocate the entire stock over time.

The analysis for the monopoly allocation is very similar. The monopolist seeks to maximize profits given by

$$\pi = \int_0^\infty \{P(Q(t),t)Q(t) - CQ(t)\}e^{-rt}dt$$

subject to

(5) $$\int_0^\infty Q(t)dt \leq S_0$$

where:

$Q(t)$ = monopolist's total production at time t

$P(Q(t,),t)Q(t)$ = total monopoly revenue at time t

Let

$$MR(Q(t),t)) = \frac{\partial}{\partial Q(t)}[P(Q(t),t)Q(t)]$$

$$= P(Q,t)[1 + \frac{1}{\varepsilon(Q,t)}]$$

where

$$\varepsilon(Q,t) = \frac{P(Q,t)}{Q(t)}[\frac{\partial P(Q,t)}{\partial Q}]^{-1}$$

= own elasticity of demand for the resource.

The monopolist's analogue of equation (4) is:

$$(6) \qquad \frac{\dot{MR}}{MR - C} = r$$

That is, the monopolist will control production from the resource stock so that his net monopoly rent rises at the market rate of interest.

The magnitude of the monopoly distortion depends on the following parameters.

a. The magnitude of extraction costs, and the variation of extraction costs with output and depletion of the resource stock.

b. The magnitude and time dependence of the demand elasticity.

c. The market discount rate and imperfections in capital markets.

With constant extraction costs and constant or increasing demand elasticity, the monopolist will be more conservation minded than competitors. This is illustrated in Figure 2.

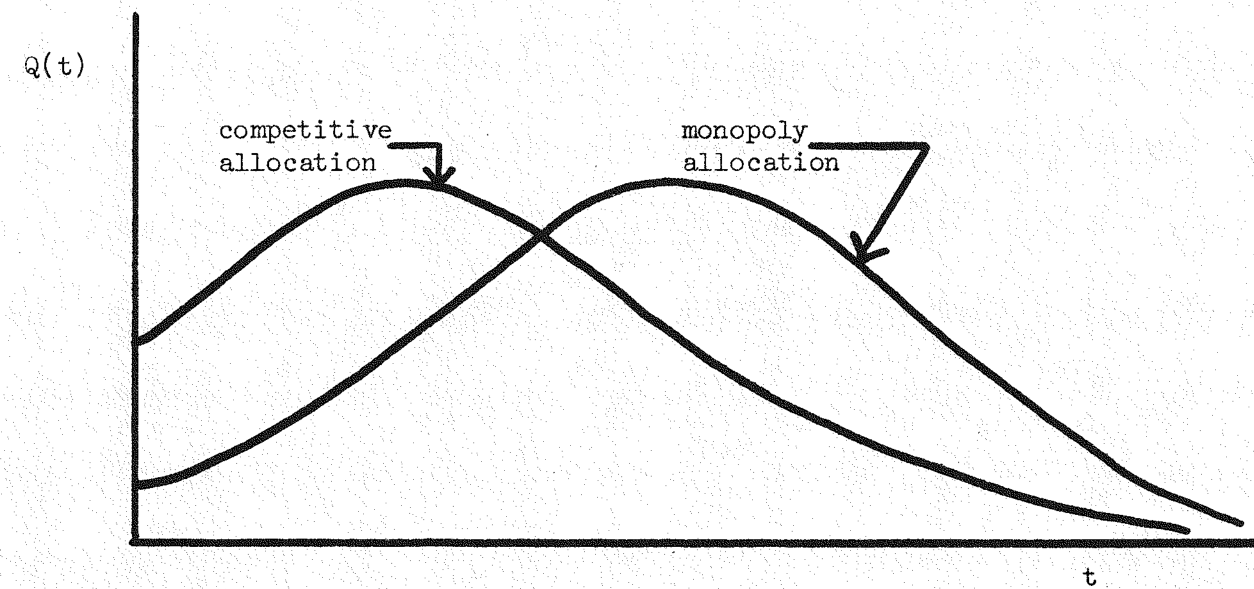


Figure 2

The project will quantitatively evaluate the distortion due to monopoly ownership of resource stocks as a function of the economic parameters specified above. Assessment of distortions for mixed cases of imperfect competition may be more difficult, but recent results of Salant [1975] and Dasgupta and Stiglitz [1975] suggest that quantitative judgments may be possible here as well.

The issue of market control in the substitutes for exhaustible resources raises several interesting questions. For example:

(*i*) Given monopoly control of an exhaustible resource, how does the monopolist's profit-maximizing extraction policy change with the introduction of substitute sources of supply?

(*ii*) Are there reasons for thinking that a monopolist has an incentive to delay the introduction of substitute sources of supply?

(*iii*) Can one say anything precise about the timing of the introduction of an innovation under differing market structures?

To place these questions in a more concrete perspective, we may consider several issues of current concern in the economics of energy resources.

(*i*) The incentives bearing on the OPEC cartel to alter production in response to innovations in the market for substitute energy sources. A related issue is the expected change in market prices upon introduction of the substitute, and the effects of the price changes on the substitute industry.

(*ii*) The possible consequences of the acquisition of coal and uranium reserves and productive capacity by major petroleum companies. Would a vertically integrated energy company with monopoly power tend to delay the introduction of coal and nuclear energy as substitutes for petroleum resources? [With due conscience, we should comment that the example we offer is intended to be a hypothetical motive for the analysis. It is not a presumption of monopolistic practices in energy markets.]

(*iii*) The implication of alternative arrangements for subsidizing research on shale oil extraction technology, such as direct government financed research versus patent rights on shale technology, on the expected time at which shale oil would be introduced as a substitute technology for natural crude oil.

It is well known (see, e.g., Nordhaus [1974] and Stiglitz [1975f]) that under perfectly competitive conditions, producers will schedule production in order to exhaust supply when the resource price equals the price of a substitute. Dasgupta and Stiglitz [1975] have extended the theory to consider the effects of the supply conditions of a substitute on the profit-maximizing extraction path of a resource monopolist. The substitute limits the monopolist's market power by placing an upper bound on the price of the resource. As in the static case, under most conditions the monopolist will extract a given stock slower than would be socially desirable. Of course, this implies that the monopolist will extract the resource at a faster rate relative to the social optimum at future times. This conservative bias delays the introduction of a substitute source. Note that in the absence of set-up costs, there is no basis for instituting price floors to encourage the introduction of a substitute. The substitute should be introduced when its price equals the resource price determined by the monopolist, and it is in the interest of the monopolist to exhaust his stock at that time.

Listed below are the four cases of polar competition in the supply of exhaustible resources and substitutes.

Ownership Pattern

Resource	Substitute Product
Competition	Competition
Monopolist	Monopolist
Competition	Monopolist
Monopolist	Competition

Dasgupta and Stiglitz [1975] have developed a methodology to analyze the qualitative effects of alternative institutional arrangements on:

- The time at which a substitute source of supply is introduced;
- The price of the substitute source of supply;
- The resource extraction path and price trajectory;
- The effect of uncertainty in the introduction of a substitute on the resource extraction path.

Consider the allocation of an exhaustible resource with zero extraction costs and an inexhaustible substitute with production costs $\bar{p}$ under alternative institutional arrangements. Thus we might consider oil as the exhaustible resource and the breeder reactor as an approximation to the substitute. Figure 3 gives allocation paths assuming different institutional arrangements. The methodology for determining these paths will be discussed in the next interim report.

Curve (a) is the socially optimal extraction policy. The resource is exhausted at time T_S and the substitute is marketed at price $\bar{p}$. Curve (b) corresponds to the case in which a monopolist owns both the resource stock and the substitute (e.g., infinite patent rights on the substitute). The monopolist delays introduction of the substitute until time T_m, and the price of the substitute is $\bar{p}(1 + \frac{1}{\varepsilon})$ where ε is the elasticity of demand. Curve (c) is the mixed case of monopoly ownership of the substitute and competitive ownership of the resource stock. This case is interesting because it sets the stage for an analysis of decentralized search for a substitute product with infinite patent rights.

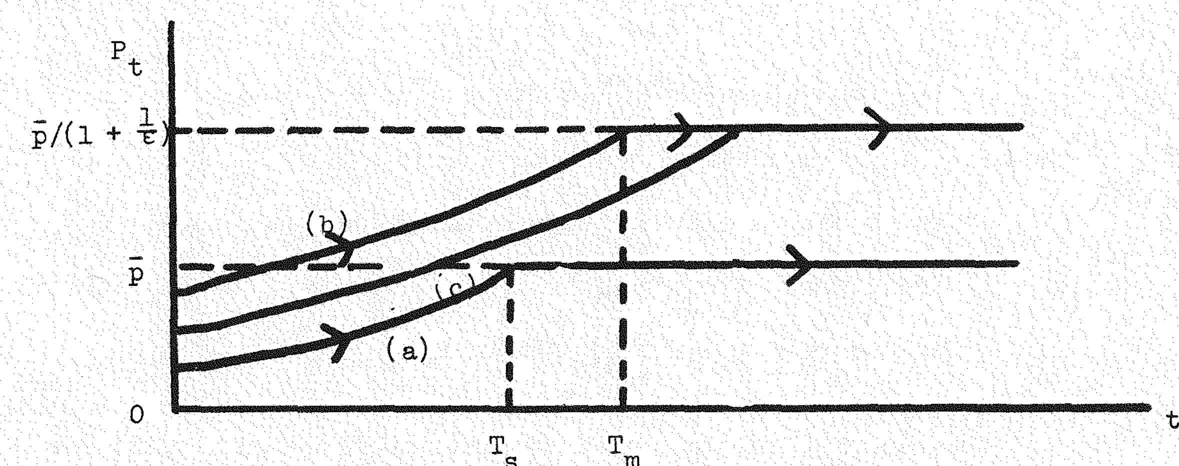


Figure 3 Resource allocation paths under alternative institutional arrangements.

It would appear that there now exists a theoretical matrix for the analysis of polar cases of competition in the supply of natural resources, on which the parameterizations required for quantitative estimates may be developed. To date, however, there has been little theoretical analysis of imperfect competition in the demand for natural resources, and the consequences of oligopolistic competition in energy markets.

III. Technical Appendix

1. Taxation and Intertemporal Bias in the Allocation of Natural Resources

1.1 Introduction

In order to analyze the distortions associated with the intertemporal allocation of a natural resource, such as oil, it is useful to divide the exploration and production process into a number of stages:

1. Pre-bidding exploration
2. Bidding and the award of leases
3. Exploration of leases
4. Development of leases (extraction)

This distinction is useful because different kinds of market imperfections are likely to impinge at different points in this exploration-development process. Some tax provisions, for instance, may primarily affect the pattern of exploration, others the pattern of development of known fields.

It is also useful to distinguish among three possible types of bias:

1. The rate of extraction from a particular well
2. The rate of extraction from a particular field
3. The rate of extraction from all fields (the rate of world or national consumption)

These distinctions are important for a number of reasons. First, some tax provisions (or other market distortions) might determine which wells or fields are used, but not affect the overall rate of consumption of the resource. Such would be the case, for instance, if all oil fields were identical, there were zero extraction costs, and the government imposed a constant percentage depletion allowance for a particular set of fields, which would be removed after a period of years. Prior to the introduction of such a tax provision, firms would have been indifferent with respect to which field they developed. The effect of this provision is to induce them to extract from this particular set of fields. But the rate of consumption of oil is unaffected.

Secondly, in our earlier discussion, we distinguished between static efficiency and dynamic (or intertemporal) efficiency. In fact, just as there are a number of facets of static efficiency, so are there a number of facets of dynamic efficiency. The one on which almost all of the literature has focused may properly be referred to as consumption efficiency. When all oil fields are identical it is the only kind of efficiency that we need be concerned with. A necessary condition for consumption efficiency is that the shadow price of the resource (in the competitive economy, the market price) rise at the rate of interest (the marginal rate of transformation between output today and output tomorrow).[1,2]

Thus, this kind of inefficiency is concerned only with the rate at which oil is consumed, in the aggregate. But there are other possible sources of inefficiency. The rate at which oil is extracted from a particular well may affect the total amount which can be extracted and the cost of extracting it; as we discussed above, there may be inefficiencies

---

[1] This is only a necessary condition; in addition, we require that the so-called transversality conditions be satisfied; in this context, this implies that either the stock of the resource be used up in the limit as $t \to \infty$. or that, if there is a produced substitute for the commodity, the stock is used up at the moment the price equals the price of the produced substitute.

[2] This is precisely true only if there were zero extraction costs. With positive extraction costs it is the difference between market price and extraction costs which must rise at the rate of interest.

in the pattern of extraction, so oil which is more expensive to extract is extracted before oil which is less expensive to extract. These inefficiencies are associated not so much with the aggregate rate of consumption, but with which fields become developed and the rate at which oil is extracted from a given field or well. The cost of these inefficiencies may be a reduction in the total stock of oil, or an increase in the present discounted value of expenditures required to extract the oil. We shall refer to this kind of efficiency as extraction efficiency.

These distinctions will prove useful in developing parameterizations designed to quantify the magnitude of the intertemporal bias associated with various kinds of distortions. In this interim report, we apply the principles developed above to the analysis of the distortions arising from tax policy.

1.2 Tax Provisions Affecting Intertemporal Bias

There are a number of provisions of the tax code which have an important effect on the pattern of extraction of oil. These include:

(a) depletion allowances
(b) immediate write-off of intangible drilling expenses
(c) write-off upon abandonment of expenditures required to obtain leases
(d) special treatment of capital gains
(e) tax deductability of interest payments on indebtedness

The first two are provisions which are peculiar to oil and other minerals; the latter are provisions which affect other sectors as well, although their distortionary impact may, in some cases, be greater here than elsewhere.

There are two ways to proceed with our analysis. We could consider sequentially each of the stages in the process of exploration and extraction, and ask how these tax provisions affect each. This would provide a more systematic and in some sense more logical approach.

The approach that we take is to begin with what appears, at least potentially, to be the most important sources of bias, and those which are most easily explained. In particular, we commence our analysis with the basic issue of consumption efficiency.

1.3 Consumption Efficiency

As a first approximation, we assume that all oil has the same, and at each moment of time, constant extraction costs. The assumption is made so that we can focus on the issue of consumption efficiency, and avoid the problems associated with "extraction efficiency." The extraction costs are assumed to decline exponentially at the rate λ.

We let

p = price of oil on the market

p_s = the scarcity price of oil, the rent obtained upon sale

p_e = the cost of extraction

Then, clearly,

(1) $$p = p_s + p_e \quad .$$

Let $D(q)f(t)$ be the demand curve for oil.[1] We assume it is shifting upwards over time, at an exponential rate. We further parameterize the demand curve by assuming constant elasticity; we thus write

$$p = q^{-1/\eta} e^{mt} \qquad m > 0$$

In this parameterization, we have made two crucial assumptions: the elasticity of the demand curve is independent of time, and the demand curve shifts uniformly over time at an exponential rate. More general parameterizations, allowing the elasticity of demand to be a function of time, say increasing exponentially, would clearly be manageable. A better function than an exponential shift one might entail first an increase in the demand, then a decrease, e.g. a quadratic function.[2] The essential difficulty in determining the appropriate parameterization is that we are required to forecast what the demand curve will look like in say 50 or 100 years, and there is no a priori reason to believe either that it will shift uniformly out, or that it will shift out for a while and then shift in. At this stage of the analysis, it is difficult to tell how sensitive any results we might obtain would be to the assumed demand curve prevailing 50 or 100 years from now. The answer will require

---

[1] It should be noted that even this apparently general function form is very restrictive; it assumes that the demand curve shifts in a neutral way, i.e., uniformly, over time.

[2] These parameterizations could be formulated

$$\ln p = -(ac^{bt})\ln q + \ln f(t)$$

where $f(t)$ is some function of time, such as exponential, and

$$\ln p = -\frac{1}{\eta}\ln q + \ln(at^2 + bt + c)$$

performing some sensitivity tests within the model to be described below.

Then, with no taxation, the price and output trajectories are described by the differential equations

(2) $$\dot{p} = rp - (r + \lambda)p_e(0)e^{-\lambda t}$$

(3) $$\dot{q} = \frac{q}{\eta}[m - \frac{rp - (r + \lambda)p_e(0)e^{-\lambda t}}{p}]$$

Given the initial price, p_0, we can solve this pair of differential equations for .

(4) $$p = \psi(t;p_0)$$

(5) $$\dot{q} = \phi(t;p_0)$$

We then integrate (5) to obtain

(6) $$\int_0^{T^*} q(t)dt = \int_0^{T^*} \phi(t;p_0)dt = S_0$$

where S_0 is the initial stock of the resource, and T^* is the date at which the resource is exhausted. If there were no substitute for the resource, and the resource were essential in production, $T^* = \infty$. If on the other hand, there exists a substitute,[1] say competitively produced

---

[1]It should be clear that we have already made one essential assumption, that the substitute is a perfect one. More generally, the effect of substitute will be to reduce the demand for the resource at any level of price. In that case, the general description of demand provided above would be applicable.

at a price $Z(t)$ at date t, with $Z' < 0$, then T^* is that date at which

$$p(T^*) = Z(T^*) \tag{7}$$

Equations (4) - (7) can be solved simultaneously for $p(t)$ and $q(t)$. For simple forms of $Z(t)$ it is even possible to obtain closed form solutions. Thus, if we assume

$$Z(t) = z$$

$$p_e = 0$$

$$m = 0$$

we obtain

$$p = p_0 e^{rt} \tag{8a}$$

$$q = e^{-r\eta t} p_0^{-\eta} \tag{8b}$$

$$\int_0^{T^*} q\,dt = \frac{p_0^{-\eta}}{r\eta}(e^{-r\eta T^*} - 1) = S_0$$

$$= Z^{-\eta} + p_0^{-\eta} = S_0 r_\eta$$

so

$$p_0 = [S_0 r\eta + Z^{-\eta}]^{-\frac{1}{\eta}} \tag{8c}$$

The higher S_0, the lower p_0, and the higher Z, the higher p_0, as expected.

We now turn to the analysis of the effects of taxation. This is a more complicated question than might appear at first sight. We introduce the following notation:

t_c	corporate tax rate
t_{cg}	tax rate applicable on capital gains
d	percentage depletion allowance

First, consider an integrated exploration-production firm. It purchases a piece of land and discovers oil; it holds the oil for some period, and then sells it. Its net cash flow appears as in Table 1.[1]

Table 1

Date	Event	Non tax cash flow	Tax cash flow
0	Land Purchase	$-p_L$	
t_1	Exploration	$-c_e$	$+t_c c_e$
t_2	Extraction	$p-p_e$	$((d-1)p+p_e)t_c$
t_3	Sale of Land	0	$p_L t_{cg}$

[1] We assume after extraction the land is worthless. Other cases are easily handled.

The return from waiting, postponing extraction, is

$$\dot{p}(1 + (d - 1)t_c) - \dot{p}_e(1 - t_c)$$

and if the firm can borrow with interest deductability, the rate of return must be equal to the interest rate:

$$(9) \qquad \dot{p}(1 - t_c + dt_c) - \dot{p}_e(1 - t_c) = r[p(1 - t_c + dt_c) - p_e(1 - t_c) + p_L t_{cg}] \ .$$

But now consider a firm which, for one reason or another, must sell its oil field before extracting from it. When will this purchaser extract the oil? His cash flow is presented in Table 2.

Table 2

Event	Non-Tax Cash Flow	Tax Cash Flow
Purchase of Lease	$-v$	0
Extraction	$p - p_e$	$(dp - (p - p_e))t_c$
Sale of Land	0	$t_{cg}v$

We thus obtain

$$(9') \qquad \dot{p}(1 - t_c + dt_c) - \dot{p}_e(1 - t_c) = r[p(1 - t_c + dt_c) - p_e(1 - t_c) + vt_{cg}] \ .$$

Consider now the problem of the marginal owner. When will he sell his lease? His cash flow is presented in Table 3.

Table 3

Event	Non-Tax Cash Flow	Tax Cash Flow
Land Purchase	$-p_L$	0
Exploration	$-c_e$	$+ c_e t_c$
Lease Sale	v	$- (v - p_L)t_{cg}$

We assume he has already done the exploration. Then he sells the lease when

$$\dot{v}(1 - t_{cg}) = r[v(1 - t_{cg}) + p_L t_{cg}] \quad . \tag{10}$$

Assume that the oil was extracted the moment after sale of the lease. Then

$$v(1 - t_{cg}) = p(1 - t_c + dt_c) - p_e(1 - t_c)$$

Hence, assuming $\dot{p}_e = \dot{t}_c = \dot{d} = 0$

$$\dot{v} = \dot{p}\frac{(1 - t_c + dt_c)}{1 - t_{cg}}$$

Thus

$$\dot{p}(1 - t_c + dt_c) = r[p(1 - t_c + dt_c) - p_e(1 - t_c) + p_L t_{cg}]$$

or

$$\dot{p} = r[p - \frac{p_e(1 - t_c)}{1 + dt_c - t_c} + \frac{p_L t_{cg}}{(1 - dt_c - t_c)}] \tag{9''}$$

Note that if (9") is satisfied, provided p_L is not too high, i.e., the original owner did not overpay for the lease, the RHS of (9') is greater than that of (9"), so it would be optimal for the purchaser to extract immediately. (9) and (9") are identical, so not surprisingly, vertical integration has no effect on the rate of extraction.

By examining Tables 1 - 3 we can also see the tax consequences of vertical integration (or disintegration), of the exploration and production processes. If capital losses can be written off against ordinary income while capital gains are taxed at the special capital gains rate, then it pays to remain integrated if and only if, at $v > p_L$

$$-v(t_c - t_{cg}) + p_L t_c \leq p_L t_c$$

it always is.[1]

The qualitative implications of this analysis can be summarized as follows:

1. The corporation profits tax, without a depletion allowance and with $p_L = 0$ (or $v = 0$) has no effect on the rate of extraction.

2. The depletion allowance, with $p_e = 0$, has no effect on the rate of extraction.

3. With $p_L > 0$ (or $v > 0$), then the corporation tax increases the rate of increase of prices, and therefore implies that initial prices today are too low, i.e., there is excessive consumption of oil.

---

[1]If $v < p_L$ (The firm has overpaid for its lease), then it is replaced by t_{cg}; firms are indifferent to vertical integration. This is true if capital gains and losses are treated symmetrically.

4. The above affect is greater if capital losses can be written off ordinary income than if it can be written off against other capital gains (at the margin).

5. With $p_e > 0$, then the depletion allowance increases the rate of increase of prices, and therefore implies that initial prices today are too low, i.e., there is excessive consumption of oil.

The intuitive interpretation of these effects is straightforward.

The ability to take a tax write-off on the capital loss associated with the depletion of a natural resource (of the value of the land or lease) leads to excessively early depletion.

The argument that the depletion allowance is necessary as a substitute for depreciation is specious, if the rights to the oil (the land on which the oil is located) can be bought and sold and capital losses and gains on those transactions are taxable or tax deductible.

If there were zero costs of extraction and the capital loss on the value of land resulting from depletion were not tax deductible, the depletion allowance would not be distortionary. The reason that the depletion allowance is distortionary even if $p_L = 0$, is that with a p ositive cost of extraction, if the rent per unit oil increases exponentially at the rate of interest, the price will increase somewhat more slowly than the rate of interest, and therefore the value of the depletion allowance, which is based on the price, not the rent, increases more slowly than the rate of interest; therefore, there is an incentive to extract early.

In the above analysis, we have identified a number of parameters which are crucial for the determination of the magnitude of the intertemporal bias resulting from the tax system: the size of extraction costs relative to market price, the value of the original purchase price of the lease or land, and the effective corporation and capital gains tax rates.

With these parameters, the differential equations (9) can easily be solved numerically, and the solutions compared with the solutions for the no tax distortion case. From this a numerical estimate both of the change in the pattern of consumption of oil as well as an estimate of the social cost of this loss can be made.

We propose that the following method be used for calculating the value of the social loss. Let $U(q)$ be the social welfare obtained from consuming at rate q. Note that

$$p = U'(q)$$

price equals the marginal utility of consumption, if we normalize the marginal utility of income at unity. For a small change in the consumption profile then

$$\Delta W = \int U(q + \Delta q)e^{-rt}dt - \int U(q)e^{-rt}dt$$

$$\frac{\Delta W}{\int pqe^{-rt}} = \frac{\int p\Delta qe^{-rt}dt}{\int pqe^{-rt}dt}$$

the present discounted value of the change in the consumption.

For larger changes in consumption, we must include a modification to take account of diminishing marginal utility of consumption. Let us approximate

$$U'(q + \Delta q) = U'(q) + U''\Delta q$$

$$= p(1 - \eta\frac{\Delta q}{q})$$

where η is the elasticity of marginal utility of consumption of the resource.

Then

$$\frac{\Delta W}{\int pqe^{-rt}dt} = \frac{\int p[\Delta q - \eta\frac{(\Delta q)^2}{q}]e^{-rt}dt}{\int pqe^{-rt}dt}$$

If the natural resource is a relatively small component of national income, then it may be reasonable to estimate η in the following way. Assume an additive utility function. Then, η is the reciprical of the price elasticity. Since for oil, this is believed to be somewhat less than unity, it implies that the appropriate value of η is slightly greater than unity.

The above calculations set out the basic methodology, for calculating both the amount of the distortion and its social costs. There are, however, a number of alternative interpretations of the effects of taxation on the intertemporal allocation of natural resources, depending on the tax treatment of sales and purchases of oil. These are set out in the next section.

1.3.1 Other Interpretations of the Effects of Taxation on Intertemporal Consumption Efficiency

The basic analysis of the intertemporal bias resulting from depletion allowances assumed that:

(a) The sale of a natural resource be treated as ordinary income.

(b) The difference between the value of the land containing the natural resource and the value after the resource has been extracted can be written off. (In the analysis, we assumed it could be written off against capital gains; under certain circumstances it may be written off against ordinary income, in which case we replace t_{cg} by t_c (or t_p) the corporate (or personal) tax rate.)

(c) Because of the interest deductability provision in the tax code the appropriate opportunity cost of capital is simply the market rate of interest (see Stiglitz [1975g]).

The first two assumptions are probably descriptively correct, but represent a significant deviation from how comparable assets are treated elsewhere in the economy; the difference between the purchase price and sale price of an asset is ordinarily subjected to capital gains taxation. Thus, the sale of oil ought to be treated as the sale of a (risky) asset which previously had been purchased (i.e. the land or lease containing the oil had been purchased), and the difference between the imputed cost of purchase (if the piece of land produced 1 million gallons of oil, and the value of the lease, once the oil is extracted, is zero, then the cost of a gallon of oil is just the cost of the lease divided by by a million) and the sale price minus the cost of the sale (i.e., the cost of extraction) be subjected to capital gains taxation. (Such tax treatment would ensure symmetry of this class of assets with other asset classes; whether it is efficient to do this depends on an analysis of the general effects of the favorable treatment of capital gains.)

The effect of this kind of treatment would be to modify the equilibrium condition to

$$\dot{p}(1 - t_{cg}) - \dot{p}_e = r(p - p_e - t_{cg}(p - p_L))$$

where p_L here is the imputed cost per unit of oil of the original purchase price of the lease.

Given the kinds of parameterizations introduced earlier, it is easy to solve numerically for the price and quantity paths of consumption, and then to compare that with the pre-tax path.

Qualitative results appear to be ambiguous.

$$\frac{\partial \dot{p}}{\partial t_{cg}} = \frac{p_e + (p_L - p_e)}{(1 - t_{cg})^2}$$

if

$$\dot{p}_e = p_e = 0 \quad ,$$

then

$$\frac{\partial \dot{p}}{\partial t_{cg}} \gtrless 0 \quad \text{as} \quad p_L \gtrless p_e$$

so for small p_e, the effect of taxes is anti-conservationist, but for large p_e it is conservationist. Other cases may be analyzed similarly.

The more questionable hypothesis is whether the appropriate rate of interest is that before tax. That depends on the ability of firms to borrow (and, given the preferential treatment of capital gains, both with respect to rates and to ability to postpone the tax, whether they actually do borrow); if firms cannot borrow at the margin, then the appropriate opportunity cost of capital may be $r(1 - t_c)$. Assume that there were no special treatment of capital gains, and capital losses could be written off ordinary income; then the appropriate differential equation is

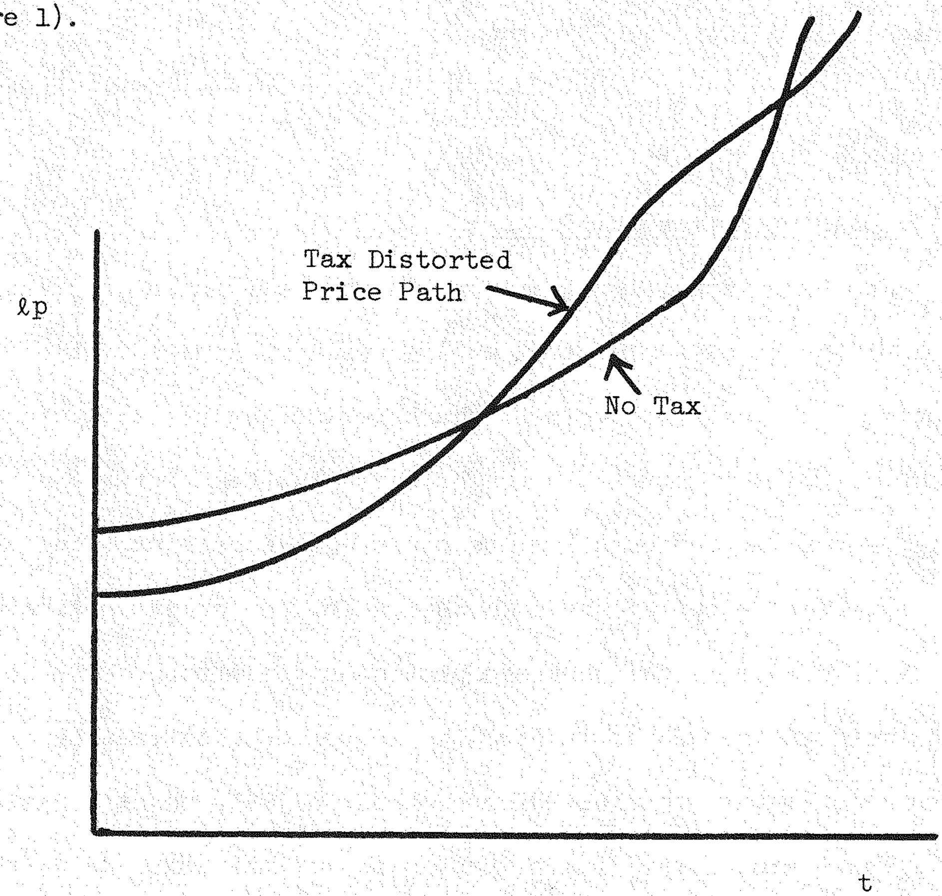
$$\frac{\dot{p} - \dot{p}_e}{p - p_e - p_L} = r(1 - t_c)$$

Then, it is unambiguously the case that the tax system leads to a too conservationist policy. The numerical solution to this differential equation proceeds as earlier.

If we combine this with an oil depletion allowance, we obtain ambiguous results:

$$\frac{\dot{p}\,\dfrac{(1 + dt_c - t_c)}{1 - t_c} - \dot{p}_e}{p\,\dfrac{(1 + dt_c - t_c)}{1 - t_c} - p_e - p_L} = r(1 - t_c) \quad .$$

If $p_e = 0 = p_L = 0$, the conservationist effect is maintained. For large p_e and d, tax policy may be anti-conservationist. As we note below, since during the growth path of the economy, p_e will initially be small and then rise, the effect of taxation may be to initially raise the price (be conservationist) and then lower its price from what it would have been, and finally, possibly, raise it from what it would have been. (Figure 1).



1.4 Distortions in the Pattern of Extraction from Known Reserves

In this section of the report, we discuss the development of parameterizations attempt to calculate the intertemporal misallocation of resources resulting from the incorrect pattern of extraction -- what we referred to earlier as extraction efficiency. We begin our analysis with a simple model, which we believe would provide a first order estimate of the magnitude of this inefficiency.

1.4.1 Constant Extraction Costs

The basic simplifying assumption we shall employ is that the amount of oil to be extracted from a given well is independent of how it is extracted, and that each well is characterized by a given, constant cost (per unit oil) of extraction. The first two tax provisions referred to earlier -- the depletion allowance and the immediate write off of drilling expenses -- can be shown to have no effect on extractive efficiency. That is, although these provisions affect the rate at which oil is extracted, they do not affect the sequence of extraction from known reserves (later, we shall discuss their affects on exploration).

There are, however, some provisions of the tax code which may have a serious effect on extractive efficiency.

Probably one of the more important sets of distortions arises from the inability to write off capital losses on an accrual basis. Thus, if a firm has paid too much for a lease, to realize the capital loss it must sell the land (lease) and let someone else extract the oil. If tax considerations were the only relevant determinants, obviously firms would simply sell the land (lease). The bias in extraction patterns then arises from the fact that on all leases for which individuals have overpaid, the purchase price of the lease by the extractor is its market value, while on all other leases, there is a random relationship between purchase price and market value, associated with the uncertainty, before exploration, of the amount of oil at any location and its cost of extraction.

As we noted in our earlier discussion, the effect of the tax on capital gains (or more accurately, the write-off of the difference in the value of the lease before extraction and after extraction), is to encourage the more rapid extraction of over-priced wells. If all wells have the same amount of oil, and differ only in extraction costs, and firms ex ante have the same priors on the probability distribution of extraction costs, then if they are risk neutral, firms will base their bids for leases on the mean value. (It is not quite accurate to say that they bid as if the probability distribution of extraction costs were concentrated at the mean, since the present discounted value of the rents from acquiring a lease are likely to be a non-linear function of extraction costs, particularly with the tax provisions being discussed.) Thus wells which turn out to have high extraction costs will be resold (and hence revalued), while those with low extraction costs will not. As is clear from equation (9) or (9"), since p_L is either a constant or declining function of extraction costs, the sequencing of extraction according to extraction costs will be correct.

Now, however, assume that there are different fields, in which individuals priors about the extraction costs differ, so that the bids for leases within the two different fields differ. It will then be the case that if the two wells turn out to have the same extraction costs, they nonetheless will differ with respect to the timing of the extraction. The well in the field with the higher prior expectation of extraction costs will be extracted after the well in the field with the lower prior expectation of extraction costs.

A second source of incorrect timing arises from provisions relating to the depletion allowance. Any binding restriction on the applicability of the depletion allowance is likely to result in some extraction efficiency. For instance, the recently passed provision leading to differential depletion allowances for large and small operators may or may not have any effect; on the one hand, it is possible that the only effect is a change in the industrial organization of the sector; the actual process of extraction (as opposed to exploration, refining, etc.) will be conducted by small operators, who can take advantage of the depletion allowance. Such a view, however, assumes that there are no clear economic advantages of vertical integration; i.e., firms that are vertically integrated are perfectly indifferent as to whether they are or are not vertically integrated. Such is not likely to be the case, so that some extraction will continue to be in the hands of the large operators. It is then clear that of two wells with identical extraction costs oil from the well controlled by the small operator will be extracted first.

Another provision of the depletion allowance is the limit to the percentage of income that can be taken as a depletion allowance. Again, this restriction may not be binding, and have its only effect on the units in to which the economic activities associated with oil extraction are organized. But if the restriction is binding, it obviously may affect the pattern of extraction, since the marginal depletion allowance may in effect be zero.

A third source of incorrect timing arises from the diligence clauses in government leases; leasees are required to extract (at least some) oil from the lease within a fixed number of years or at least face the risk of loosing the lease. This means that if a lease turns out to have oil with extraction costs which are high, but still less than the market price, oil from the well will be extracted; the firm will postpone development of the field as late as it can, but even though efficiency would require still further postponement, the oil is extracted.

A fourth source of incorrect timing arises for bidding. Our earlier analysis showed that a lease which might be economically viable, i.e. from which extraction ought to take place in the future, might be sold by the original owner, in order for him to take a capital loss, if he overbid in acquiring the lease. Restrictions in the ability to transfer title to a lease might then lead to abandonment of an economically viable well; and if the lease is not put up for releasing sufficiently soon, this may lead to postponement of extraction from the economically efficient time.

Similar effects arise from royalty bidding where royalties are a percentage of price, rather than net profits. Again, abandonment will occur whenever the ratio of extraction costs to market price exceeds one minus the royalty rate (for all dates in the future). This will have a distortionary effect if (and only if) the lease is not put up for bidding again sufficiently soon.

The effect of these (and possibly other) provisions of tax and leasing policy is that the sequencing of extraction from wells is inefficient. The economic cost of this misallocation can in principle be easily calculated; all we need to do is to compare the present discounted value of extraction costs under the correct sequencing with that under the actual pattern of extraction. The data required to make that calculation are, however, probably not available. As a first rough estimate, one might proceed as follows.

Take the distribution of observed extraction costs as being the distribution of extraction costs of all wells. Take an estimate of the amount of oil to be extracted in each of the next T_0 years (say some exponential rate of increase in the amount of oil consumed). Then, over that period, one could easily calculate the present discounted value of extraction costs if there were no sequencing, i.e., if $\bar{p}_e$ is the mean extraction cost, and $Q(t)$ the amount of oil extracted at time t, then it is just

$$\bar{p}_e \int_0^{T_0} c^{-rt} dt \quad .$$

Let $F(p_e)$ be the distribution function of wells (oil) by extraction costs. Then we can easily solve for $p_e(t)$, the extraction costs of oil extracted at time t, assuming that the total stock of oil was just

$$\int_0^{T_0} Q(t)\,dt$$

from the equation

$$p_e(t) = F^{-1}\left(\frac{\int_0^t Q(t)\,dt}{\int_0^{T_0} Q(t)\,dt}\right) .$$

We then calculate the present discount value of extraction costs as

$$\int_0^T p_e(t)Q(t)e^{-rt}dt .$$

These calculations are fairly easily performed numerically. There are three systematic sources of errors in this method. First, the present distribution of oil by extraction costs is not a good representation of the true distribution, since there is some presumption that although the economic forces leading to early extraction of cheap oil are incompletely effective, they still work to some extent. This means that our procedure overestimates the amount of oil with low extraction costs, and accordingly overestimates the inefficiency present in the market. Secondly, it assumes that the extraction costs of all oil are known and that all oil has been discovered. In fact, what needs to be compared is the cost of extraction from unknown sources. This interaction between exploration and extraction will be discussed more fully below. Again, the effect of ignoring this is probably that our procedure overestimates the real loss. Finally, the finite horizon of our formulation probably leads to an underestimate of the resource cost of having an incorrect pattern of extraction. But with reasonable discount rates, this is not likely to be economically significant. Present research is concerned with refining these estimation procedures.

1.5 Further Work in Estimating the Tax Induced Inefficiencies in the Intertemporal Allocation of a Natural Resource.

Work presently under way extends the framework developed in the preceding analysis in two ways: first, we consider the affect of tax (and leasing policy) on the pattern of exploration, and the interaction of this with the pattern of extraction. Secondly, the preceding analysis assumed that the costs of extraction were a technological datum. In fact, there are strong interactions between the rate of extraction, the cost of extraction, and the total amount of oil extracted. We are in the process of analyzing the effects of tax and leasing policy on the choice of patterns of extraction for a given field.

References

Atkinson, A. [1970], "On the Measurement of Inequality," Journal of Theory.

Atkinson, A. [1975], The Economics of Inequality.

Arrow, K. [1964], "The Role of Securities in the Optimal Allocation of Risk Bearing," Review of Economic Studies, April, pages 91-96.

Arrow, K. and Kurz, M. [1970], Public Investment, the Rate of Return, and Optimal Fiscal Policy, Johns Hopkins Press.

Arrow, K. [1973], "Rawls Principle of Just Savings," Swedish Journal of Economics.

Cass, D. and Yaari, M. [1966], "A Note on the Role of Money in Providing Sufficient Intermediation," American Economic Review.

Cass, D. and Stiglitz, J. [1970], "The Structure of Investor Preferences and Asset Returns and Separability in Portfolio Allocation," Journal of Economic Theory, June.

Dasgupta, P. [1973], "Some Theoretical Explorations in the Economics of Exhaustible Resources," Bayerisch Academy Symposium in Systems Analysis, Reisenburg.

Dasgupta, P. [1974], "On Some Alternative Criteria for Justice between Generations," Journal of Public Economics.

Dasgupta, P. and Heal, G. [1974], "The Optimal Depletion of Exhaustible Resources," Review of Economic Studies, Symposium, pages 3-28.

Dasgupta, P. and Stiglitz, J. [1975], "Uncertainty and the Rate of Extraction under Alternative Institutional Arrangements," Stanford University mimeo.

Debreu, G. [1959], The Theory of Value.

Diamond, P. [1967], "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty," American Economic Review, 57, No. 4, September, pages 759-776.

Fellner, W. (ed.) [1967], Ten Essays in Honor of Irving Fisher.

Gilbert, R. [1975], "Issues in Resource Depletion under Uncertainty," Stanford University mimeo.

Green, J. [1973], "Information, Efficiency and Equilibrium," HIER Discussion Paper No. 284, March.

Grossman, S. [1975], "The Existence of Futures Markets, Noisy Rational Expectations and Informational Externalities," IMSSS Technical Report No. 182, Stanford University.

Hahn, F. [1966], "Equilibrium Dynamics with Heterogeneous Capital Goods," Quarterly Journal of Economics.

Harsanyi, . [1975], "Can the Maximin Principle Serve as a Basis for Morality," American Political Science Review, June, pages 594-607.

__________ [1975], "Cardinal Welfare, Individual Ethics, and Interpersonal Comparisons of Utility," in Economic Justice, E. Phelps (ed.)

Hart, O. [1974], "On the Optimality of Equilibrium when Markets are Incomplete," mimeo.

Heal, G. [1975], "Interaction between Resource and Financial Markets," unpublished manuscript.

Hirshleifer, J. [1971], "The Private and Social Value of Information and the Reward to Inventive Activity," American Economic Review, Volume 61, No. 4, pages 561-574.

Hotelling, H. [1931], "The Economics of Exhaustible Resources," Journal of Political Economy, 39, April, pages 137-175.

Kay, J. and Mirrlees, J. [1975], "The Desirability of Natural Resource Depletion," Oxford University mimeo.

Kolm, S. [1972], Justice and Equity.

Koopmans, T. [1960], "Stationary Ordinal Utility and Impatience," Econometrica, 28, pages 287-309.

Leland, H. [1974], "Production Theory and the Stock Market," Bell Journal of Economics and Management Science, Volume 5, Spring.

Malinvaud, E. [1953], "Capital Accumulation and Efficient Allocation of Resources," Econometrica, 21, April, pages 233-268.

Nordhaus, W. [1974], "The Allocation of Energy Resources," in Brookings Papers on Economic Activity, . Okun and . Perry, (eds.).

Peterson, . [1975], "Two Externalities in Petroleum Exploration," in G. Brannon (ed.) Studies in Energy Policy.

Phelps, E. [1975], Economic Justice.

Ramsey, E. [1928], "A Mathematical Theory of Savings," Economic Journal, December.

Rawls, J. [1971], A Theory of Justice.

Rothschild, M. and Stiglitz, J. [1975], "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," IMSSS Technical Report No. 170, Stanford University.

Salant, S. [1975], "Nash-Cournot Equilibrium for an Exhaustible Resource Like Oil," unpublished manuscript.

Samuelson, P. [1954], "The Pure Theory of Public Expenditure," Review of Economics and Statistics.

__________ [1958], "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money," Journal of Political Economy, LXVI, No. 6, pages 467-482.

Shell, K., Sidrausky, M. and Stiglitz, J. [1967], "Capital Gains, Income and Saving," Review of Economic Studies, January, pages 15-26.

Shell, K. and Stiglitz, J. [1967], "Allocation of Investment in a Dynamic Economy," Quarterly Journal of Economics, November, pages 592-609.

Solow, R. [1974], "Intergenerational Equity and Exhaustible Resources," Review of Economic Studies, Symposium, pages 29-45.

__________ [1975], "The Optimal Extraction of Natural Resources," unpublished manuscript.

Spence, M. [1973], "Job Market Signalling," Quarterly Journal of Economics August, pages 355-379.

Stiglitz, J. [1972a], "Optimality of Stock Market Allocation of Investment," Quarterly Journal of Economics, 86, No. 1, February, pages 25-60.

__________ [1972b], "Portfolio Allocation with Many Risky Assets," in Mathematical Methods in Investment and Finance, K. Shell and G. Szego, (eds.).

Stiglitz, J. [1974a], "Growth with Exhaustible Resources: The Competitive Economy," Review of Economic Studies, Symposium, pages 139-152.

__________ [1974b], "Growth with Exhaustible Resources: Efficient and Optimal Growth Paths," in Review of Economic Studies, Symposium, pages 123-139.

__________ [1974c], "On the Non-Optimality of Stock Market Allocations," Stanford University mimeo.

__________ [1975a], "Simple Formulae for the Measurement of Inequality," Stanford University mimeo.

__________ [1975b], "Monopolistic Competition and the Capital Market," IMSSS Technical Report No. 161, Stanford University.

__________ [1975c], "The Theory of Screening, Education and the Distribution of Income," American Economic Review, June.

__________ [1975d], "Tax Policy and the Oil Industry," in G. Brannon (ed.) Studies in Energy Policy.

__________ [1975e], "Information and Capital Markets," Oxford University mimeo.

__________ [1975f], "Monopoly and the Rate of Extraction of Exhaustible Resources," American Economic Review, (forthcoming).

__________ [1975g], "The Corporation Tax," Journal of Public Economics, 5, pp. 303-11.

Sweeney, J. [1974], "Economics of Depletable Resources: Market Forces and Intertemporal Bias," Stanford University mimeo.

Yaari, M. [1965], "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," Review of Economic Studies.

2. The Effect of Tax Concessions on Energy Conservation

Most energy conservation measures require the substitution of capital for current energy use. Better insulated buildings cost more but have lower energy consumption, machines with higher thermal efficiencies are typically more sophisticated and more expensive, and integrated steel mills which economize on reheating steel have higher capital costs. In many cases the initial investment decision severely restricts the extent to which energy can be subsequently economized without completely redesigning or replacing the capital stock, and it is therefore important to ask whether individuals will choose the socially efficient degree of energy conservation when they make their investment decision, or whether there will be a systematic bias towards underconservation.

A generous explanation for the existence of corporate tax concessions is that they are either designed to mitigate the distortionary impact of the fiscal system, or to counteract systematic misperceptions on the part of decision makers. Granting tax exemptions to bond interest payments and depreciation fall in the first category, whilst the infant industry argument for tariff protection often rests on the second explanation.

Both considerations appear to have motivated the suggestion that investments in energy conservation should attract preferential tax treatment. This paper examines the impact of four concessionary schemes on investment in energy conservation, namely

(i) Accelerated Depreciation. The IRS code of 1954[9] permitted U.S. businesses to depart from straight-line depreciation and use one of two new accelerated methods -- the double-declining balance (ddb) or the sum-of-years-digits (syd) methods. In 1971 a further liberalization allowed firms to combine these in the most favorable way. In continuous time the <u>syd</u> method is always more attractive, so the 1971 reform can be thought of as minimizing the errors introduced by making the calculations on an annual basis. We therefore confine our attention to the continuous time variant of <u>syd</u>. The firm is allowed to deduct D_t on an asset whose original cost K was incurred t years previously, where

$$D_t = \frac{2K}{\ell^2}(\ell - t) \quad , \quad 0 \leq t \leq \ell \quad . \tag{1}$$

The present discounted value of depreciation allowances at the date of installation is then

$$K\psi(\ell) = \frac{2K}{\ell^2}\int_0^\ell (\ell - t)e^{-rt}dt \quad ,$$

or

$$\psi(\ell) = \frac{2}{r\ell}\left(1 - \frac{(1 - e^{-r\ell})}{r\ell}\right) \tag{2}$$

In this formula ℓ is the lifetime over which the asset can be depreciated, and r is the rate at which the company discounts after-tax profits. Table 1 below gives selected values of ψ for varying combinations of r and ℓ.

<u>Table 1</u>

Value of Depreciation Allowances

$r\ell$	0.0	0.1	0.25	0.5	0.7	0.9	1.0	1.3	1.5	1.75	2.0	2.5	3.0
$\psi(\ell)$	1.0	0.97	0.92	0.85	0.82	0.76	0.74	0.68	0.64	0.60	0.56	0.51	0.45

Clearly, the shorter the allowable lifetime ℓ the more valuable is the tax concession, whilst, conversely the higher the money rate of interest, that is, the higher the rate of inflation, the less valuable it will be. "Accelerating" depreciation by reducing ℓ can therefore be used to mitigate the distortionary effects of inflation on the current U.S. fiscal system, as Shoven and Bulow [1975] demonstrate.

(ii) Investment Tax Credit. This instrument allows a firm to immediately write off a fraction δ of the original cost against taxes without affecting the pattern of depreciation allowances. The fraction allowable may depend on the asset life, and is currently 10% for assets with a life of more than 7 years, 6 2/3% for assets with a life between 5 and 7 years, and 3 1/3% for those with a life between 3 and 5 years.

The effect of the tax credit is similar to an accelerated depreciation allowance, further concentrating the deductions early in the life of the equipment. We shall be interested in the sum $\delta + \tau\psi(\ell)$, the present value of the tax concessions when τ is the corporate tax rate.

(iii) Tax Exempt Bonds. The IRS might allow firms to finance a fraction α of certain categories of investment (such as energy conservation) by issuing tax exempt bonds paying a rate of interest ρ and having a maturity m. This allows the firm to raise capital at a lower cost than the normal bond rate, R, for if the marginal rate of tax on those tax payers who are indifferent to holding tax exempt and nonexempt bonds is γ, then

(3) $$\rho = (1 - \gamma)R$$

(iv) Loan Guarantees. The government might offer loan guarantees for certain categories of investment, under which the government would meet any potential bankruptcy liabilities which arise purely because of these investment expenditures. Loan guarantees are particularly suited to cases in which the government wishes to offset pessimism on the part of firms about the value of investments. Since the government exercises considerable control over the price of energy (via tariffs, taxes, price ceilings, concessions, regulation, rationing etc.) the government is likely to have better information about the future, and to be in a position to offer insurance or reassurance via the loan guarantee. In what follows we will assume that such guarantees are not needed, or have already been provided, for we ignore the effects of uncertainty. It is possible that the loan guarantee will give the company a higher credit rating on bond issue, and the effect is somewhat similar to allowing it to issue some fraction of its debt in tax exempt bonds. The concession will thus not be separately distinguished in what follows.

The Effect on Energy Conservation

The case for special tax treatment of investment in energy conservation measure must rest on either the belief that firms systematically misperceive the benefits of conservation, or that they fail to allow for externalities. The most obvious externality is pollution, gaseous, particulate, or thermal, air or water born. Most of these depend on the type of fuel and combustion technology as well as the quantity of fuel burnt. Subsidizing the reduction of energy consumed may be a very inefficient method of reducing emmissions, for it will provide absolutely no incentive to install any devices which reduce emissions per unit of energy consumed, such as cooling towers, scrubbing stacks, catalytic after burners, etc., nor will it encourage the optimal choice of fuel. Indeed, subsidizing equipment designed to reduce pollution emissions is totally ineffective in the absence of a tax on emissions, for there will still be some cost to installing the devices, and no benefit to the installer.

There is a completely different kind of externality which may provide a case for the subsidization of energy conservation, and that arises where future supplies of and/or demands for energy are uncertain. Current energy consumption decisions may then affect the uncertainty facing other decision makers, and may alter their decisions. Preliminary investigations by Stiglitz suggest that competitive markets do not internalize this externality, but it is far from clear what kind of intervention is required to improve allocative decisions. In the absence of a well articulated framework within which to handle this kind of externality it seems desirable to concentrate on the effects of misperception as opposed to externalities.

The natural model to use is a vintage capital model, and in earlier work (Newbery [1976]) such a model was used to show that in an otherwise efficient environment (i.e. no taxes or externalities) if firms underestimated the rate of energy price rise they would underinvest in energy investments. Here, however, we are interested in modelling the choice of investment in the presence of potentially distorting taxes, and will therefore use a slightly different model. For convenience, the main results of the earlier paper are summarized in the appendix.

An investment K will reduce energy consumption by an amount $E = F(F)$. As far as the profits of the firm are concerned, it is as though the investment produced energy E for sale. The equipment lasts T years without impairment of efficiency, at which point it becomes worthless. The price of energy is p_t at date t.

The investment is eligible for accelerated depreciation over a lifetime ℓ, investment tax credit of fraction δ, and can finance a fraction α with tax exempt bonds paying ρ, of maturity m. The profits tax rate is $\tau \equiv 1 - \theta$.

The first question to be asked is whether the firm will choose to finance the investment out of retained earnings or by issuing bonds (ignoring, for the moment the case of tax exempt bonds). If R is the bond rate and r the rate at which the company discounts profits available for retention or distribution (i.e. profits after corporation tax), then the choice is between retention (with a cost of \$1) or a bond reducing retentions by \$$\theta R$ each year for m years, with repayment of \$1 after m years. The present cost of the latter is

$$\theta R\left(\frac{1 - e^{-rm}}{r}\right) + e^{-rm} = \frac{\theta R}{r}(1 - e^{-rm}) + e^{rm} \quad .$$

If $\theta R > r$, then this will be larger, and bonds will never be issued. Since we observe that firms do choose to issue bonds, and also finance investment out of retained profits, it seems that $r = \theta R$. This would be a convenient assumption, since the timing and maturity of bond issues is then irrelevant, and we shall make it, but it is important to draw attention to the strength of this assumption. If it is not true (and there is a variety of reasons and empirical observations why it may not be true) then we need a more complex model of firm behavior than is currently available. The question of what determines the cost of capital is complex and is here largely assumed away.

Given this simplification the firm is indifferent between finance out of retentions and bond finance, so the present value of investment is

(4a) $$V = -K + \alpha\theta K\int_0^m (R - \rho)e^{-rt}dt$$

(4b) $$+\theta F(K)\int_0^T p_t e^{-rt}dt$$

(4c) $$+(\delta + \tau\psi(\ell))K \quad .$$

Line (4a) is the present value of issuing tax exempt bonds αK and reducing non tax exempt bonds by αK, leaving the initial cost K to be financed by K bonds or retained earnings. Line (4b) gives net profits after paying tax at rate τ on the imputed sales of energy over the lifetime of equipment T, whilst line (4c) is the present value of the tax saved through the depreciation allowances.

If the price of energy is expected to rise at a steady rate g (in terms of money), and if $\rho = (1 - \gamma)R$, $r = \theta R$, then

$$V = \theta F(K)p\phi(r - g) - K\{1 - (\delta + \tau\psi) - \alpha\gamma(1 - e^{-rm})\}$$

where

$$\phi(x) \equiv \frac{1 - e^{-xt}}{x}$$

and p is the current price of energy. This is maximized with respect to K when

$$(5) \qquad F_K p\phi(r - g) = \frac{1 - (\delta + \tau\psi) - \alpha\gamma(1 - e^{-rm})}{\theta} .$$

The optimal choice of investment is given by

$$(6) \qquad F_K^* p^* \phi(r^* - g^*) = 1$$

where stars denote efficient prices (or, in the case of g^*, correct forecasts of future efficient prices). We can now examine the extent to which allowances are needed to offset tax distortions (in particular, θ) and misperceptions (particularly divergencies between g and g^*).

1. <u>No Misperceptions</u>

If $r = r^*$, $g = g^*$, $p = p^*$, then the correct choice of technique requires the RHS of equation (5) to be unity. As is well known, this will be achieved if:

(i) $\alpha\gamma = 0$ i.e. no tax exempt bond issues <u>and</u>

(ii) $\delta = 0$, $\ell = 0$ i.e. no tax credit and instantaneous write off ($\psi(0) = 1$).

In this case efficiency will be preserved as the rate of inflation varies, providing interest rates adjust.[10/] If $\ell > 0$, then either δ or α must be a function of ℓ, and the system will not be inflation proof.

<u>Misperceptions about Future Energy Prices</u>

To correct for misperceptions about p and/or g, taxes, etc. must be arranged so that

(7) $$1 - (\tau\psi(\ell) + \delta) - \alpha\gamma(1 - e^{-rm}) = \frac{p\theta\phi(r - g)}{p^*\phi(r^* - g^*)} .$$

It seems reasonable to assume that θ, m and γ are given, which implies that any of α, δ, or ℓ can be varied to achieve equality in equation (7). If firms perceive general inflation, but not the increasing <u>relative</u> price of energy, then the right hand side of (7) will be independent of the rate of inflation. Equation (7) can be rearranged thus

(8) $$\alpha\gamma e^{-rm} - (1 - \theta)\psi(\ell) = \frac{p\theta\phi(r - g)}{p^*\phi(r^* - g^*)} + \alpha\gamma + \delta - 1$$

where the RHS is independent of the rate of inflation. The LHS can also be made independent of inflation by setting $\alpha = 0$ and $\ell = 0$, that is, granting instantaneous depreciation and not allowing tax exempt bond issues. If, for some reason ℓ and m cannot be altered, then α should be set to minimize the sensitivity of the LHS to variations in r, whilst δ is adjusted to ensure average equality of equation (8). That is

(9) $$\alpha = \frac{\tau e^{rm}}{\gamma m} \frac{\partial\psi}{\partial r} .$$

Thus if $\tau = 48\%$, $m = 15$ years, $\ell = 12$ years, $r = 8\ 1/2\%$, $\gamma = 30\%$, then $\alpha = 0.95$. If $m = 20$ years, $\ell = 30$ years, $\alpha = 1.45$. (Such members correspond to present U.S. parameters for equipment and structures respectively.) On this basis setting $\alpha = 1$ could be justified as minimizing the effects of inflation on tax law (assuming, that is, that the obvious reform of instantaneous depreciation is precluded).

The Effects of Price Control

If, for some reason, the government regulates the price of energy so that $p_t < p_t^*$, then the simplest method of inducing the correct choice of investment is to set $\alpha = \ell = 0$

$$\delta_t = 1 - \theta\{1 - \frac{p_t\phi(r - g)}{p_t^*\phi(r^* - g^*)}\}$$

For example, if $p_0 = p_0^*$ at some date in the past, and if $g^* = r^* = r$, and $p_t = p_0$ (i.e. prices are frozen) then

$$\delta_t = 1 - \theta(\frac{1 - e^{-rT}}{rT})e^{-gt}$$

If $r = 5\%$, $T = 15$ years, $\theta = 1 - \tau = 52\%$, then $\delta_0 = 63.4\%$, $\delta_{10} = 77.8\%$. Of course, if $\ell > 0$, (that is, if immediate tax write-offs are not allowed) then the tax credit δ will have to be increased by $\tau(1 - \psi(\ell))$ and will be even larger than these already large figures.

The implications of price control can be summarized as follows. If the ratio of market price to social opportunity cost (p/p^*) is kept constant, and if firms foresee the rising price correctly $(g = g^*)$ then an extra tax credit

$$\delta = 1 - \theta(1 - \frac{p}{p^*})$$

is all that is required. If producers are misled about g (as they are likely to be under price control) then the tax credit (or other subsidy for conservation) will have to be larger, whilst if the market price of energy is frozen whilst the cost rises, then the subsidy δ will have to change through time. The problems of altering the tax system to achieve adequate conservation with price control are formidable, and constitute yet another argument against price control.

Divergences between Social and Private Discount Rates

The most obvious source of distortion is the tax system which drives a wedge between the rate of return on investment and the intertemporal rate of substitution. With an interest income tax the rate of return on capital will differ from the rate of intertemporal substitution, and full optimality cannot be achieved. It is a moot point whether r is above or below r^*. If $R \simeq r^*$, that is, the rate of investment were optimal, and were bond financed at the margin, then $r < r^*$, offsetting the tendancy to underinvest in energy conservation. Without further work on the relationships between the social discount rate, the bond rate, and the discount rate used by firms it is difficult to be sure of the direction of the bias. However, if empirical work is able to resolve these issues the formulae set out above should prove helpful in any restructuring of the tax system.

Conclusion

The best treatment of energy conservation investments is for the government to publish best available forecasts about future prices of energy, to allow instantaneous depreciation of investments, and to bring the firm's discount rate into line with the social discount rate. Failing this, tax exempt bonds have the merit of reducing the sensitivity of other allowances to the rate of inflation, and, coupled with the best choice of initial tax credit, in principle allow the government to make any desired correction to misperceptions about future energy prices and inappropriate depreciation allowances.

Footnotes

1/ The analysis performed in section 3 of the second interim report suggests that the current OPEC price exceeds the equilibrium price for a perfectly competitive world petroleum market.

2/ If consumption is too conservative, the market may not signal long-run disequilibrium if there are no substitutes and an infinite planning horizon. The difficulty is that excess supply is not observed at any finite time.

3/ Forward markets would be of value in the allocation of risk when demand and supply conditions are uncertain.

4/ Of course variations in price due to changes in information will occur and are not impeded by the existence of forward markets. These changes occur simultaneously with the revelation of new information, and do not exploit price variations when forward markets are complete.

5/ It may pay one country to stabilize its internal price provided the monopolist still finds it profitable to continue his random price strategy.

6/ We are assuming here that the alternative energy source would be produced at a price below the price set by the cartel, and in sufficiently large quantities to pose a threat to cartel profits.

7/ Alternatively, an optimal tariff program will diminish monopoly profits and increase consumer welfare (see section 3.3). Such a program may be pursued independently of policies to promote research and development.

8/ It may be argued that it is better to "buy American" and the way around the dilemma described above is to restrict imports. However, the factors required to produce the alternative energy source have uses elsewhere in the economy, and the opportunity costs to society cannot be neglected on the basis of domestic versus foreign production. Unemployment and foreign exchange limitations cannot be ignored, but energy policy is certainly not the most effective instrument for dealing with these problems. At best, one might justify a "Shadow cost" that lowers somewhat the cost of domestic energy relative to imports, but there are few persuasive arguments for a substantial cost differential.

9/ The details are culled from Shoven and Bulow [1975].

10/ This is a big proviso. Feldstein, Green and Sheshinski [1976] analyzed the distortionary impact of inflation on the current tax system. We are implicitly assuming an inelastic supply of domestic savings.

3. Project Independence and Optimal Trade-Offs in Flexibility

The dramatic increase in the price of Arab crude-oil in 1973 gave rise to the idea of Project Independence. Unfortunately it was not made clear precisely what Independence means. It is occasionally interpreted as a move towards autarky. But the cost of autarky is high and it is clearly questionable whether such a move is optimal. Moreover, even the demand for autarky is left vague since the timing of such a move is left unclear (e.g. should there be a great deal of dependence on foreign oil today and more tomorrow or should there be a little dependence on both these dates?) Do policies need to be enacted which in some sense"protects" the importing country from possible "embargoes," and if so, what is not clear is the nature of such policies? In what follows we analyze this question by supposing the availability of some simple policy instruments, such as border tariffs and domestic extraction policies.

But before we do this, there are some general points to be discussed. The sudden price rise of Arab crude in 1973 is usually interpreted as being a consequence of a recognition on the part of OPEC that it can exercise its monopoly power. If this interpretation is correct then the price rise could be regarded as a once for all increase due to the formation of the cartel. The movement of the price of an exhaustible resource and the price of its extraction in a world where there is a dominant extractor with a competitive fringe was analyzed in detail in our second interim report. But usually when the fear of possible embargoes is raised it is not to this foregoing feature that attention is drawn. Rather, what is being alluded to is the possibility of occasional threats to the availability of foreign supply arising out of the formation of the cartel. It is this latter possibility that gave rise to the demand for Project Independence and it is this that we shall discuss in this section.

It is obvious that if we adopt a policy of national energy independence, then the prices charged by OPEC (or an embargo by OPEC) would have no domestic consequences. The cost of such total energy autarky is high, and such a policy has few serious proponents. We are thus concerned here with partial independence. There are major categories of policies designed to increase our independence: the first, reduces our total demand for energy, and hence our need for imports; the second increases our domestic supply of oil. The former set of policies makes us worse off (if the embargo or price rise never occurs), because we consume less oil than we otherwise would; there is some oil, whose immediate benefit exceeds its cost, which is not being imported because, were we to import it now, we would be worse off if it were withdrawn.

The trade-offs in the second case are somewhat more subtle: if we produce more oil today, we will have less oil in the future; and thus, if we increase our flexibility today by producing more, we decrease our flexibility in the future. This intertemporal trade-off in flexibility seems to have gone largely unnoticed in the popular literature. It is important: there is reason to believe that, purely on grounds of flexibility, we consumed an excessive amount of oil in the 50's; although the oil quotas were justified in terms of the desirability of independence from foreign oil, in fact they have lead to a higher level of long run dependence. In addition to these two categories there is a third: without changing levels of demand and supply today, we could change the elasticity of the demand or supply curve. In the following sub-sections we propose to analyze each of these two flexibility measures in detail.

3.3 Flexibility and Intertemporal Supply Allocation

To analyze the optimal intertemporal allocation of oil we assume that there is a stock of oil to be consumed this period or next

$$\bar{Q} = Q_1 + Q_2$$

where $\bar{Q}$ is the total (domestic) stock of oil, Q_1 is extracted this period, Q_2 is extracted next. The oil has constant extraction costs, c; this period either we can purchase an arbitrarily large amount of oil from the foreign cartel at a price p_1 or it will impose a total embargo. The price next period will either be p_2 (again with essentially a horizontal supply schedule in the relevant region) or again the cartel will impose a total embargo (this simplifies the analysis; a partial embargo may also be easily analyzed). The contingencies of the two embargoes are assumed to be independent, occuring with probabilities Π_1 and Π_2. We assume the country must decide before it knows whether there will be an embargo about its extraction policy this period (and hence how much is to be extracted next period).

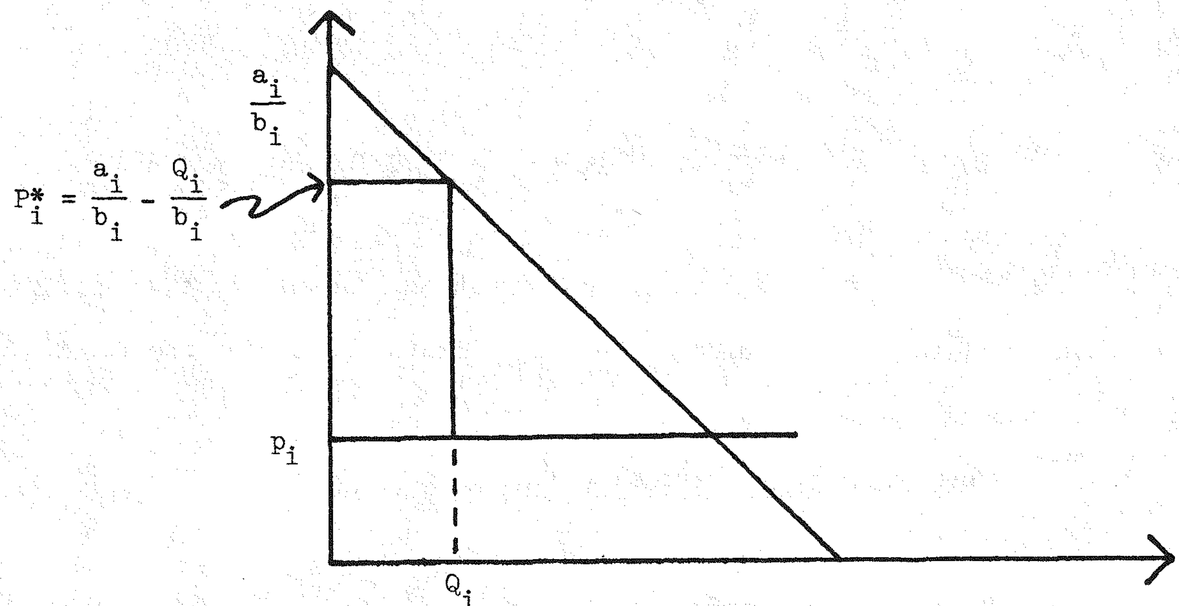


Figure 1

Again for simplicity, we assume linear demand function

$$Q_i^c = a_i - b_i p_i$$

where Q_i^c is the amount consumed the ith period. Social welfare is measured as the sum of consumer plus producer surplus:

$$\max \theta = \Pi_1[(\frac{a_1}{b_1})Q_1 - cQ_1] + (1 - \Pi_2)[(p_1 - c)Q_1 + \frac{(\frac{a_1}{b_1} - p_1)(a_1 - b_1 p_1)}{2}]$$

$$+ \frac{\Pi_2}{1 + r}[(\frac{a_2}{b_2} - \frac{Q_2}{2b_2})Q_2 - cQ_2] + (1 - \Pi_2)(p_2 - c)Q_2 + \frac{(\frac{a_2}{b_2} - p_2)(a_2 - b_2 p_2)}{2}$$

s.t.

$$Q_1 + Q_2 = \bar{Q}$$

$$(1) \qquad \Pi_1[(\frac{a_1}{b_1} - \frac{Q_1}{b_1}) - c + (1 - \Pi_1)(p_1 - c)$$

$$= \frac{\Pi_2}{1 + r}[(\frac{a_2}{b_2} - \frac{Q_2}{b_2}) - c] + (1 - \Pi_2)\frac{(p_2 - c)}{1 + r} \quad .$$

With $\Pi_1 = \Pi_2 = 0$, we obtain the conventional result that we extract all oil the first period or second as rents are higher the first period or second -- $(p_1 - c_1) \gtrless p_2 - c_2/1 + r$.

More generally, we can rewrite (1) as

$$\bar{p}_1 - c_1 = \frac{(\bar{p}_2 - c_2)}{1 + r}$$

$$p_i = \Pi_i p_i^* + (1 - \Pi_i)p_i$$

$$p_i^* = \frac{a_i}{b_i} - \frac{Q_i}{b_i}$$

where $\bar{p}_i$ is the expected price the ith period. Thus the expected price rises at the rate of interest given the expected confirmation that with correct anticipation of the probabilities, a competitive market will provide the correct intertemporal allocation of "flexibility."

The implications for the intertemporal allocation of oil are straightforward. Let

$$\varepsilon_i = -\frac{\partial \ln Q_i^c}{\ln p_i} = \frac{b_i p_i}{Q_i} \text{ , the elasticity of demand,}$$

and

$$s_i = \frac{Q_i}{Q_i^c} \text{ , the share of oil supplied domestically.}$$

Then

$$(2) \qquad \frac{\Pi_1 p_1}{\varepsilon_1}(1 - s_1) - \frac{\Pi_2 p_2(1 - s_2)}{\varepsilon_2(1 + r)} = \frac{p_2 - c_2}{1 + r} - p_1 - c_1 \quad .$$

The right hand side represents the marginal cost of flexibility, the cost of shifting a unit of production of oil from this period to next, provided the embargo does not occur. Assume for instance, as in much of the discussions, that $c_i \approx 0$, costs of extraction are approximately zero, and price is rising at the rate of interest. Then (2) implies that

$$\frac{1 - s_1}{1 - s_2} = \frac{\varepsilon_1}{\Pi_1} / \frac{\varepsilon_2}{\Pi_2}$$

import shares should be proportional to demand elasticities and inversely proportional to the probability of an embargo.

If elasticities are roughly constant, and $\Pi_1 = \Pi_2$, this implies constant import shares. Notice that without costs associated with flexibility, a slight change in the rate of increase of oil price will lead all oil to be extracted this period or next, when the value of flexibility is taken into account, a slight change in the rate of increase of prices will have only a slight effect on the intertemporal pattern of extraction.

Monopoly

Assume now that the domestic oil supply is monopolized. The monopolist is risk neutral, correctly perceives the probability of an embargo, and maximizes expected rents, i.e.

$$\text{Max } \Pi_1(R(Q_1) - cQ_1) + \frac{\Pi_2}{1 + r}(R(Q_2) - cQ_2)$$

$$+ (1 - \Pi_1)(p_1 - c_1)Q_1 + (1 - \Pi_2)\frac{Q_2(p_2 - c_2)}{1 + r}$$

i.e.

$$(R_1' - c_1)\Pi_1 - \frac{(R_1' - c_2)\Pi_2}{1 + r} = (1 - \Pi_2)\frac{p_2 - c_2}{1 + r} - (1 - \Pi_1)(p_1 - c_1)$$

where

$$R_i' = p_i(1 - \frac{1}{\varepsilon_i})$$

i.e.

$$\bar{p}_1 - c_1 - \frac{\bar{p}_2 - c_2}{1 + r} = p_1 \frac{\Pi_1}{\varepsilon_1} - \frac{\Pi_2 p_2}{\varepsilon_2(1 + r)} \quad .$$

Again, note that if $\Pi_1 = \Pi_2$, i.e. equal probability of an embargo, $\varepsilon_1 = \varepsilon_2$, $c_1 = c_2 = 0$, and the "normal" price is rising at the rate of interest then the monopolist allocates oil intertemporally in the same way that a competitor would.

But if for instance there is a declining elasticity of demand $(\varepsilon_2 < \varepsilon_1)$ or $\Pi_2 > \Pi_1$ (the probability of an embargo is increasing) then

$$\bar{p}_1 - c < \frac{\bar{p}_2 - c_2}{1 + r}$$

implying a lower price this period, in the event of an embargo, than is socially desirable; this in turn implies a higher rate of consumption. Conversely if $\varepsilon_2 > \varepsilon_1$, or $p_2/(1 + r) < p_1$. There is some presumption that the price is rising more slowly than the rate of interest (because of positive extraction costs) but that the elasticity of demand may be increasing (because of the development of substitutes). These two effects offset each other, and hence there is no clear presumption on whether a monopolist would be excessively conservationist or not.

Conservation and Demand

Suppose that the long run foreign supply curve for the resource in question is perfectly elastic at the price $\bar{p}$. This is the price in normal years. We assume it constant because, say, it is the limit pricing behavior of this cartel. (See our second interim report.) But we assume that foreign suppliers (the cartel) apply occasional shocks and raise the price by a given fraction δ. Such interruptions are assumed to occur with probability Π. It is interesting to consider the socially optimal domestic policy under such circumstances. For commodities such as oil that are inputs in production and not direct consumption goods it is important to distinguish between the long run and short-run derived demand functions. Assume for the moment that domestic production of the resource is nil, so that domestic demand is met entirely by imports. Let x denote the final consumer good which yields gross consumer surplus of $\$u(x)$. x is produced by the resource flow, Q, and a stock of capital, K, via the production function

(1) $$x = x(Q,K) \quad .$$

The rental on capital is fixed at r, and p is the random foreign price of the resource. The accompanying diagram denotes the long-run derived demand curve for the resource as AB. This is the schedule that describes the demand Q for the resource at different prices for the resource when the capital stock can be chosen optimally as well. It is computed from the solution of the problem

(2) $$\max_{Q,K} [u(x(Q,K)) - pQ - rK] \quad .$$

Notice that in (2) both Q and K are being chosen optimally for every realization of the foreign price.

In the short-run K cannot be varied. Suppose, nevertheless, that Q can be chosen. For a given value of K (say $\bar{K}$) there is a short-run demand for the resource given by the solution of the problem.A.1/

(3) $$\max_{Q} [u(x(Q,\bar{K})) - pQ - r\bar{K}]$$

A typical short-run derived demand curve is drawn as A'B'. For different values of $\bar{K}$ we have different short-run derived demand curves. The optimal policy is then obtained from solving the following problem:

(4) $$\max_{\bar{K}} E[\max_{Q} (u(x(Q,\bar{K})) - pQ - r\bar{K}) \quad ,$$

where E is the expectation operator. Problem (4) tells us that Q and $\bar{K}$ must be so chosen that

(5) $$u'(x)\frac{\partial x}{\partial Q} = p$$

and

(6) $$E(u'(x(Q,\bar{K}))\frac{\partial x}{\partial \bar{K}}) = r \quad .$$

Equation (5) and (6) tell us the capacity, $\bar{K}$, that needs to be installed. That is, the equations describe the short-run derived demand curve the economy ought to be operating on. It is of course immediate that provided the private sector assesses the probability of embargoes in the same manner as the government does the solution given by equations (5) and (6) is achieved in a competitive market. Thus, for example, if A'B' (see figure 1) is the short-run derived demand curve associated with the optimum capacity, then during normal years Q_1 will be imported (at the price $\bar{p}$), whilst in embargo years Q_2 will be imported (at the price $(1 + \delta)\bar{p}$).

Note that by the envelope theorem,

$$\frac{dU}{dp} = -Q(K,p)$$

i.e. the short run loss in utility from an increase in the price from p^* to p is

$$\Delta U(t) = \int_{p^*}^{p} Q(K,p)dp \quad .$$

The loss in utility changes (is reduced) as K adjusts, eventually attaining its long run value. Hence the total loss in utility is

$$\int_0^{\infty} \Delta U(t)e^{-rt}dt \quad .$$

The transitional loss in utility -- the excess of loss for the incorrect choice of K is given by

$$\int_0^{\infty} \Delta U(t)e^{rt} - \int_{p^*}^{p} Q(K^*(p),p)dp \quad .$$

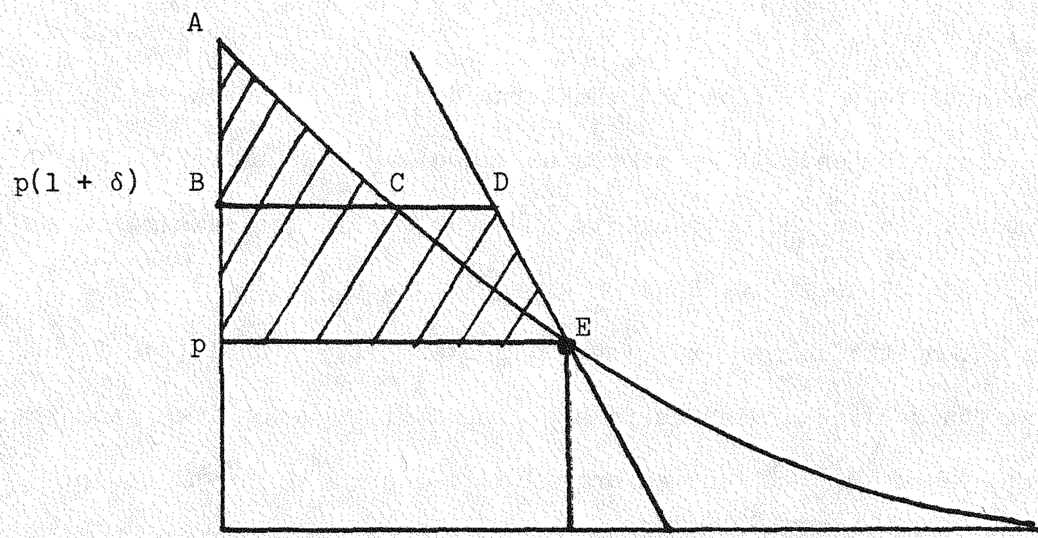


Figure 2.

Diagramatically, the extra loss in consumer surplus is the area between the short run and long run demand curves -- areas CDE. The net consumer surplus where "p" had been planned on and $p(1 + \delta)$ occurs is the difference between the areas ABC and CDE.

To obtain some insight into the structure of the problem it will be useful to parameterize somewhat. We begin with the case where $u(x)$ is iso-elastic and the production function $u(Q,K)$ is of the Cobb-Douglas form. To be specific, suppose

$$(7) \qquad u(x) = \frac{x^{1-v}}{1-v} \,, \quad v > 0 \text{ and } v \neq 1$$

$$u(x) = \log x \qquad (\text{for this case } v = 1)$$

and

$$x = Q^{\alpha} K^{1-\alpha} \,.$$

The long run derived demand curve for the resource can now be easily computed. Since (2) represents the optimization exercise Q and K are chosen so that

$$(8) \qquad \alpha Q^{[\alpha(1-v)-1]} K^{(1-\alpha)(1-v)} = p$$

and

$$(1 - \alpha) Q^{\alpha(1-v)} K^{-(\alpha+v(1-\alpha))} = r \quad .$$

Eliminating K from the two equations in (8) yields the equation

$$(9) \qquad Q^{-\frac{v}{(\alpha+v(1-\alpha))}} = \frac{p}{\alpha}\left(\frac{r}{1-\alpha}\right)^{\frac{(1-\alpha)(1-v)}{\alpha+v(1-\alpha)}}$$

In other words the absolute value of the elasticity of derived demand for the resource is $\alpha + v(1 - \alpha)/v$. It is independent of the price.[A.2/]

The short run derived demand curve is equally simple to compute. Using (7) in equation (5) yields, as above

$$(10) \qquad \alpha Q^{[\alpha(1-v)-1]} \bar{K}^{(1-\alpha)(1-v)} = p \quad ,$$

where $\bar{K}$ is obtained from equation (6). The absolute value of the short run elasticity of derived demand is $1/[1 - \alpha(1 - v)]$. This too is independent of the price.

We are here concerned with the optimal policy as expressed in (4) on the assumption that (7) is an appropriate parameterization. If Q_1 is the optimal import when $\bar{p}$ is the price and Q_2 the optimal import when $(1 + \delta)\bar{p}$ is the price, equations (5) and (6) can be expressed as

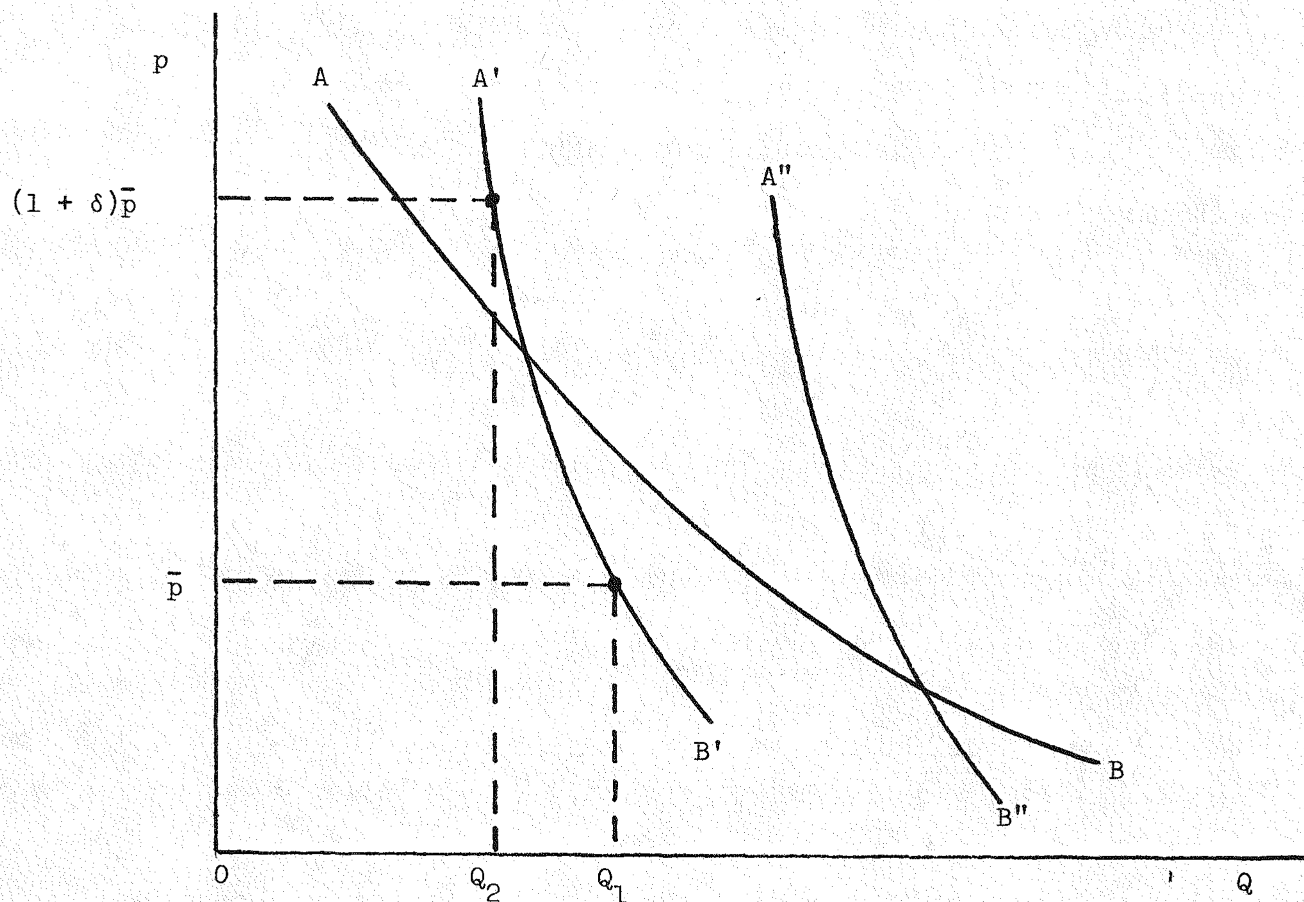


Figure 3.

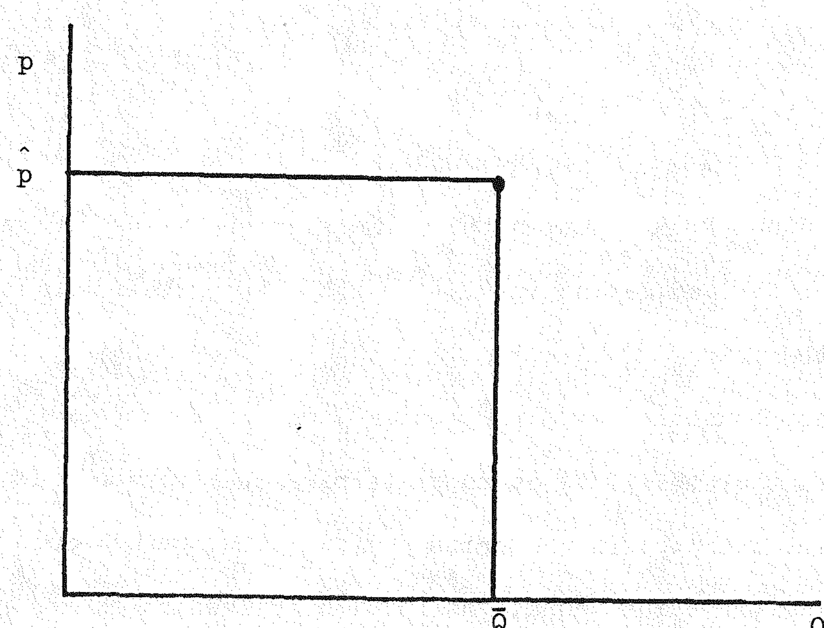


Figure 4.

(11) $$\alpha v Q_1^{-[(1-\alpha)+\alpha(1+v)]} K^{-v(1-\alpha)} = \bar{p}$$

(12) $$\alpha v Q_2^{-[(1-\alpha)+\alpha(1+v)]} K^{-v(1-\alpha)} = (1 + \delta)\bar{p}$$

and

(13) $$(1 - \Pi)v(1 - \alpha)[Q_1^{\alpha}K^{1-\alpha}]^{-(1+v)} Q_1^{\alpha}K^{-\alpha} + v\Pi[Q_2^{\alpha}K^{1-\alpha}]^{-(1+v)}(1 - \alpha)Q_2^{\alpha}K^{-\alpha} = r$$

There are three unknowns, Q_1, Q_2 and K and three equations (11) - (13). The interesting fact though is that on dividing equation (12) by (11) one has

(14) $$\left(\frac{Q_1}{Q_2}\right)^{(1-\alpha)+\alpha(1+v)} = 1 + \delta \quad ,$$

and therefore, the ratio of the optimal quantities to be imported in the two states of nature (i.e. the two possible foreign price levels) is independent of Π (i.e. the probability of the embargo occurring). Of course, the absolute import levels do depend on Π, but not the ratio. It is readily seen that this result, the fact that under (7) the ratios of the optimal import levels are independent if the probability assessment holds true irrespective of the number of possible price levels the cartel may choose from.

The parameterization assumed in (7) supposes unitary elasticity of substitutions between fixed capital (K) and resource flow (Q). A special case of considerable importance is not this but one where the elasticity of substitution is nil. That is, where there are fixed coefficients in the production of x. Thus $x = \min(Q,K)$. A typical

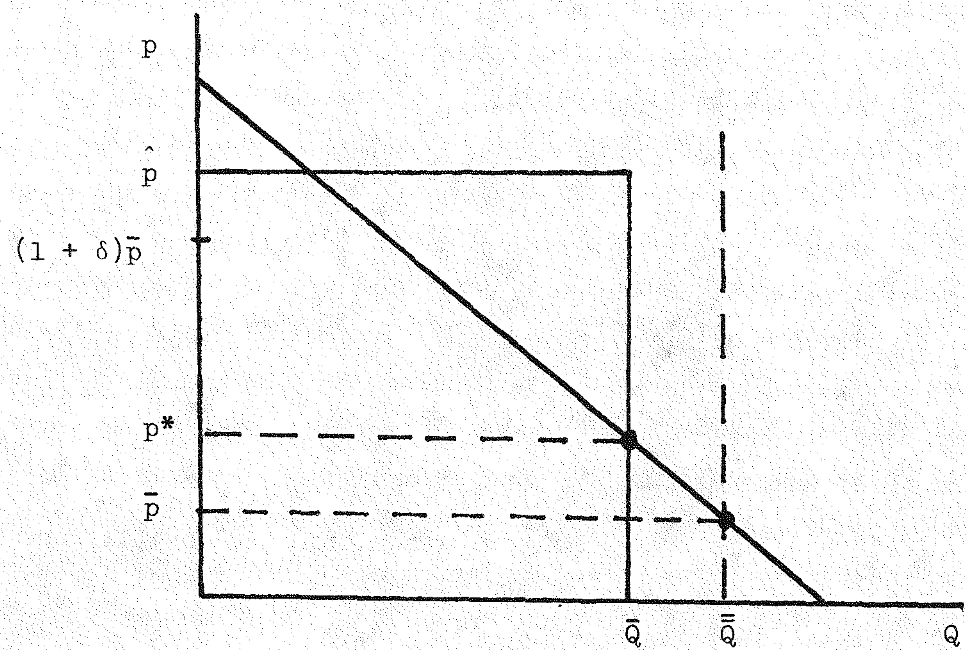


Figure 5.

short-run derived demand curve for Q then has the form given in Figure 2. That is, for a given capacity $\bar{K}$ the quantity $\bar{Q}$ is demanded inelastically up to a price $\hat{p}$ at which point demand gets choked off to zero.

Suppose that the long-run derived demand curve for the resource is linear, and takes on the form

$$p = a - bQ \quad . \tag{7}$$

Figure 3 brings together the long and short run demand curves for the resource for this special case. If there is no threat of embargo at all then the optimum policy (which is achieved under competitive conditions) is to install capacity at a level $\bar{\bar{k}}$, so that $\bar{\bar{Q}}$ is the resource requirement at each period (see figure 3). If the embargo does not occur, and Q is the quantity imported, then net consumer surplus is

$$S = \left(\frac{a}{b} - \frac{Q}{2b}\right)Q - \bar{p}Q \quad .$$

If the price rises to $(1 + \delta)\bar{p}$ the economy loses an amount in increased expenditure on imports of $\delta\bar{p}Q$. It follows that the fixed quantity that ought to be imported is one that maximizes $S - \Pi\delta\bar{p}Q$. (Π is the probability of the embargo). This is readily calculated to be at the level $\bar{Q}$ given by

(8) $$a - b\bar{Q} - \bar{p} = \Pi\delta\bar{p}$$

When there is no embargo (i.e. during normal years) the domestic price of the resource is $p^* = a - b\bar{Q}$. Thus

(9) $$p^* - \bar{p} = \Pi\delta\bar{p} \quad .$$

Thus the domestic marginal value exceeds the import price by an amount which is the product of the amount of the price rise $(\delta\bar{p})$ and the probability of the price rise occuring (Π). If the private sector incorrectly believes that $\Pi = 0$, then $(p^* - \bar{p})$ represents the tariff that should be imposed on the import and the resource. If $\delta = 1$ and $\Pi = .2$ and individuals are completely mypoic then the optimum tariff rate $(p^* - \bar{p})/\bar{p}$ is obtained as 20%.

Flexibility with Price Control

If the market price is allowed to freely adjust and if consumers hold correct expectations about the risks of embargo then no tariff is needed. If, however, the government does not allow the domestic price to rise during the embargo then even if consumers hold correct expectations they will choose an inadequate amount of flexibility unless a tariff is imposed.

Suppose that price control operates as follows. In normal years import price is $\bar{p}_t$, the tariff is τ_t, and the domestic producer price is $\bar{p}(1 + \tau)$. Producers supply a constant amount Q_p at this price (which, to simplify the analysis, we shall suppose is invariant to any <u>constant</u> tariff. This will hold if, for example, extraction costs are zero everywhere). In embargo years, the producer price is not allowed to rise, and consumer prices are set at the average price p_c

$$p_c = \frac{(\bar{Q} - Q_p)(1 + \delta)\bar{p} + Q_p(1 + \tau)\bar{p}}{\bar{Q}}$$

With correct expectations consumers choose $\bar{Q}$ so that $a - b\bar{Q} = Ep = (1 - \Pi)\bar{p}(1 + t) + \Pi p_c$. The optimal tariff is such that $Ep = \bar{p}(1 + \Pi\delta)$ from equation (9), so that

$$1 + \Pi\delta = (1 - \Pi)(1 + \tau) + \Pi(1 - S)(1 + \delta) + S(1 + \tau)\Pi$$

or

$$1 + \tau = \frac{1 - \Pi(1 - S(1 + \delta))}{1 - \Pi(1 - S)}$$

where $S = Qp/\bar{Q}$, the share of domestic production in total supply. If $S = 1/2$, $\Pi = 0.2$, $\delta = 1$, then the optimum tariff is $\tau = 1/9$ or 11%, instead of the 20% in the myopic non price control case discussed above.

<u>More General Measures of Flexibility</u>

In the previous sections we have been concerned with the costs associated with the sudden imposition of an embargo or an increase in the foreign supply price. In the analysis, it is important not to confuse the deliterious effects of a <u>high</u> price of oil from the deterious effects of a change in the price. Schematically, we can imagine the process of adjustment to a change in the foreign price of oil as in figure

In the short run, there is a greater loss of consumer surplus from having the wrong technology, and a greater "transfer" payment to the foreign producer, both because of the short run supply inelasticity makes domestic supply smaller and the short run demand inelasticity makes domestic demand larger than after full adjustment.

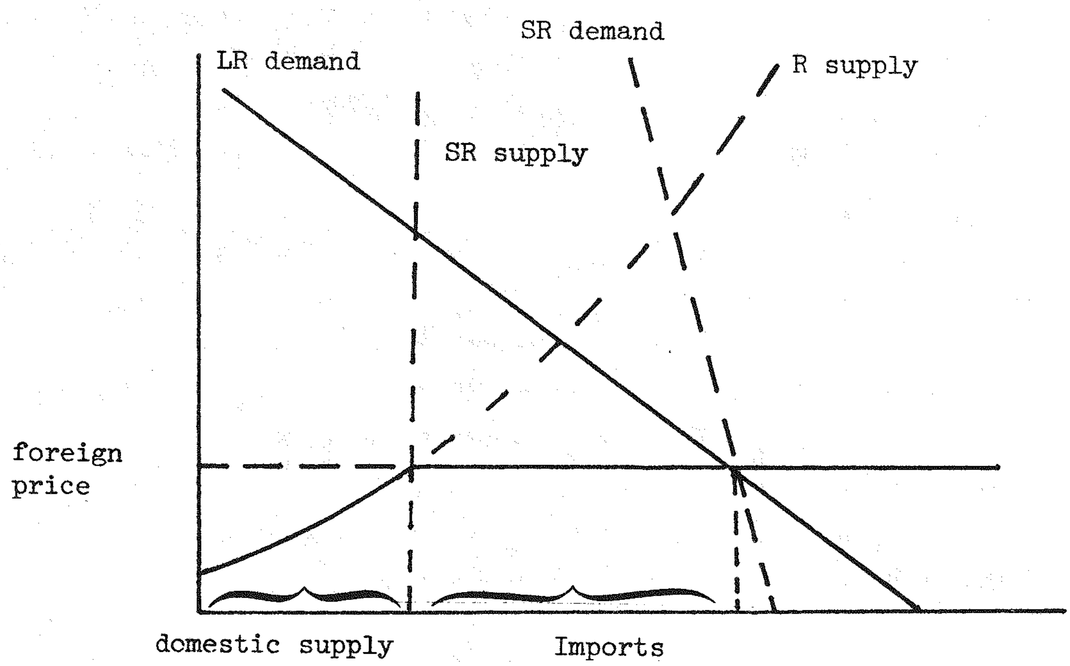


Figure 6.

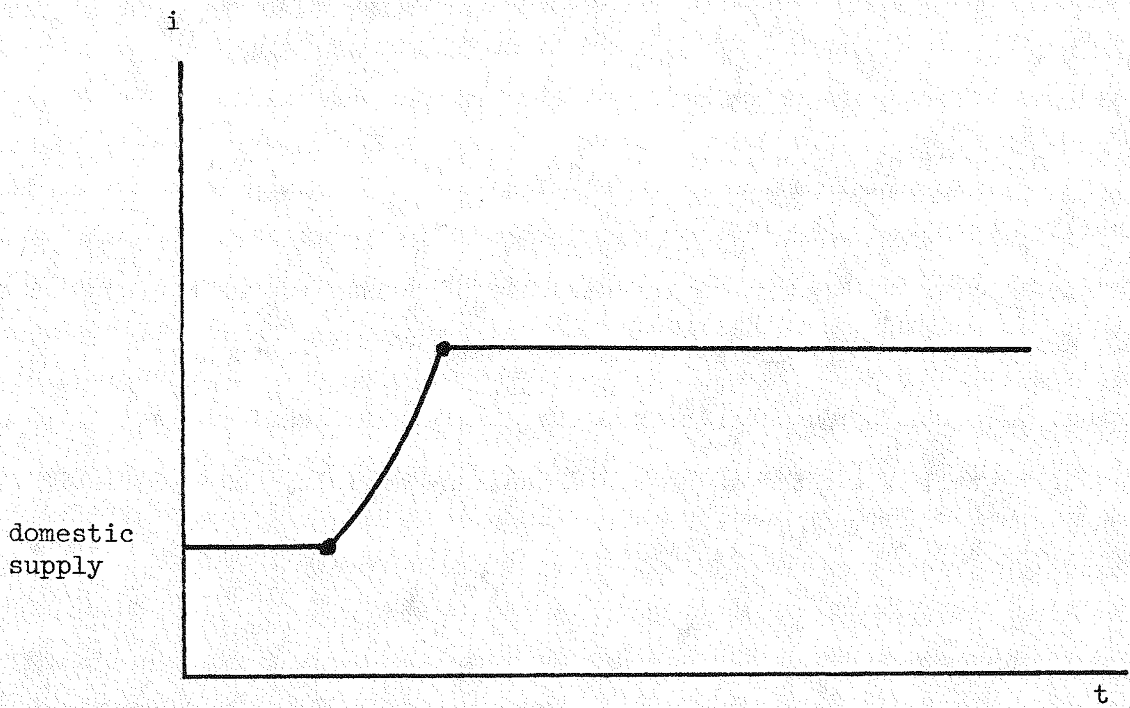


Figure 7

Returning to equation (5) we can write the level of consumption as a function of K and p:

$$Q = Q_c(K_c, p) \quad .$$

Substituting into the utility function, consumer surplus can be written as

$$u(x(Q(K_c, p)K) - pQ(K_j p) - r(K = S_c(K_c, p) \quad .$$

Domestic producer (supply) surplus can similarly be written as

$$Q_s = Q_s(K_s, p)$$

$$S_p(K_s, p)$$

when K_s is the producer capital stock. The level of K_s and K_c depends on the long run value of p and the change in p (in foreign supply).

Thus, we let $K_s = K_s(t)$, $K_c = K_c(t)$, and Q_I be the value of imports. Hence

$$Q_s(K_s(t), p(t)) + Q_I = Q_c(K_c(t),p(t))$$

from which we solve for $p(t)$. Substituting into the expressions for consumer and producer surplus, we obtain

$$S \equiv \int S_p(K_s(t),p(t))e^{-rt} + \int S_c(k_c(t),p(t)e^{-rt}dt \quad .$$

Simple parameterization may easily be developed. The value of flexibility will be associated with the speed of adjustment of K_c and K_s, the consumer's and producer's short run demand function to then be a demand function, and the magnitude of the difference between the short run and long run functions (reflecting the importance of fixed factors in consumption and production.

4. Energy Conservation with Durable Equipment and Imperfect Foresight

Here we ignore the tax system and concentrate on imperfect foresight. If investors fail to perceive that the price of energy must rise (at least when the rental component of the price is significant), then they will choose equipment which is wasteful of energy, and as a result incur two social costs -- that of having inappropriate captial, and, indirectly, by raising the demand for energy they will tend to raise its current price (which will mitigate the tendency to waste) and, if there are non-zero extraction costs, to raise the rate of price increases, which will worsen the inefficiency in the choice of investment.

Consider a vintage model in which the range of substitution possibilities at the design stage is represented by the production function

$$Y = AK^aE^bL^c \qquad a + b + c = 1$$

where Y is output, K capital stock, E is current energy use, and L is the other current input labor. Once chosen E and L cannot be varied over the fixed lifetime T of the equipment. The expected price of energy relative to output is p_t at date t, of labor is $w_t = w_0eg^t$, and the real rate of interest is constant at r. The investor chooses K,E,L to minimize the present discounted cost of producing $\bar{Y}$:

$$\min_{K,E,L} \; K + E\int_0^T p_t e^{-rt}dt + Lw_0\int_0^T e^{(g-r)t}dt$$

s.t.

$$AK^aE^bL^c \geq \bar{Y} \quad .$$

If $p_t = p_0 e^{\alpha t}$, then

$$\frac{E}{E^*} = \frac{\phi(r - g)}{\phi(r^* - g^*)} \left\{ \frac{\phi(r^* - \alpha^*)}{\phi(r - \alpha)} \right\}^{1-b} \quad , \quad \phi(x) \equiv \frac{1 - e^{-xT}}{x} \quad ,$$

where E^* is the efficient level of energy use, associated with correct forecasts and interest rates, indicated by stars. This formula shows that if the only source of bias is a low forecast for α the energy use will be excessive. It is also true that E will rise as r falls, and, to a first approximation, if rT is not too large,

$$\frac{E}{E^*} = 1 + \frac{1 + b}{4}(r^* - r)T + \frac{g - g^*}{2}T + \frac{1 - b}{2}(\alpha^* - \alpha)T \quad .$$

Underestimates of g will tend to lower E, somewhat mitigating the effects of myopic forecasting.

Correcting the Bias Towards Energy Waste

If the government were convinced that all investors held the same, wrong, estimates of α, and if all equipment were equally durable, then the optimum tax on energy is readily calculated to be at an *ad valorem* rate τ where

$$1 + \tau = \frac{\phi(r^* - \alpha^*)}{\phi(r - \alpha)} \quad .$$

However, if different equipment has different lifetimes then no single tax on energy will suffice though the correct choice of K/E can be induced by a suitable combination of taxes on equipment which decrease with durability, and a uniform tax on energy. Thus the ad valorem tax on capital would, on the assumption made above, be τ_k where

$$1 + \tau_k = \frac{T^*}{T}\left(\frac{1 - e^{-rT}}{1 - e^{-rT^*}}\right), \quad \frac{\partial \tau_k}{\partial T} < 0$$

with

$$1 + \tau = \frac{rT^*}{1 - e^{-rT^*}}$$

for some value of T^*, chosen so the distortion caused by raising production costs c would be minimized.

There are two obvious objections to this suggestion. The more fundamental is that a uniform tax on energy would not work if investors hold diverse opinions about α, and would not be necessary if they could be persuaded to hold correct expectations. If reliable information about α is available, investors will find it profitable to use it, which suggests that indicative planning is a more suitable policy to follow.

Secondly, taxes which vary with the durability of equipment can be expected to alter the choice of durability. This might not matter if biased expectations of α lead to systematic biases in the choice of durability, since we might wish to correct this distortion as well, but it appears that a lower forecast for α can either increase or reduce durability, lending no support for a tax on capital which varies with its durability. The reason for this ambiguity about the effect on the choice of durability is interesting, and worth stating.

The higher the forecast values of α, the more capital intensive will production be, and, other things being equal, the longer will it be before the capital becomes obsolete through rising energy costs. But there is some incentive to offset this tendancy to increase durability, since an early collapse of machinery will permit a change to better practice technology. The actual choice of durability will reflect a balance between these two tendancies. The next model illustrates this ambiguity.

The Choice of Durability

Suppose the production function is again Cobb-Douglas $Y = AK^aE^b$. (No other production function allows the problem to be formulated as a stationary regeneration problem, though as we are constructing an example of an ambiguous influence of α on durability this restriction should not matter.) Let the investment cost of installing capacity K with durability T be $Kg(T)$, and calculate the present cost at date t of installing and operating capacity sufficient to produce output $\bar{Y}$ for T years.

$$C_t = K_t g(T) + E_t p_t \phi(r - \alpha)$$

Now,

$$K_t = \bar{Y}k_t^b \quad , \quad E_t = \bar{Y}k_t^{-a} \quad ,$$

so

$$C_t = \bar{Y}\{k_t^b g(T) + p_t k_t^{-a}\phi(r - \alpha)\} \quad .$$

If

$$k_t = ke^{\alpha t} \quad ,$$

then

$$C_t = \bar{Y}e^{\alpha bt}\{k^b g(T) + p_0 k^{-a}\phi(r - \alpha)\}$$

and the present discounted cost of replacing the capital stock with current best produced equipment every T years is

$$C = \underset{k,T}{\text{Min}}\ \bar{Y}\frac{\{k^b g(T) + p_0 k^{-a}\phi(r - \alpha)\}}{1 - e^{-(r-b\alpha)T}}$$

whence

$$kg(T) = \frac{ap_0}{b}\phi(r - \alpha)$$

and

$$\frac{kg'(T) + p_0 e^{-(r-\alpha)T}}{kg(T) + p_0\phi(r - \alpha)} = \frac{e^{-(r-b\alpha)T}}{\phi(r - b\alpha)} \quad .$$

If $g(T) = e^{\beta T}$, which seems a reasonable durability function, then for $\alpha = 0$

$$e^{rT} = 1 + \frac{r}{\beta}$$

whilst if $\alpha^* = r$

$$(b + \beta aT)\left(\frac{e^{raT} - 1}{raT}\right) = 1$$

solves for T^*.

Then if $r = \beta = 7\%$, $T = 9.9$ and if $a = 0.5$ $T^* < T$ whilst if $a = 0.4$ $T^* > 10$. Thus $T \lesseqgtr T^*$ depending on r, β, and a.

5. Trade Policy and Import Tariff

Nearly 50% of the current annual petroleum needs of the United States are met by imports; and the U.S. is a major importer by international standards. The U.S. therefore has some monopsony power in the world oil market, and it is natural to ask what is its optimal import policy. Domestic oil producers are also large corporations and the government may also be interested in how best to tax the rental components in their profits. These two tax problems have certain similarities, but also important differences, which we discuss in this appendix.

The theoretical analysis of exhaustible resources has to date largely ignored the "geo-political realities" which so concern policy makers, and has instead concentrated on the logically prior problem of analyzing market equilibrium in an autarkic economy (which might be the whole world). This framework is suitable for the study of competitive equilibrium, and, more debatably, monopolistic equilibrium, but is ill-suited to analyze the taxation of those exhaustible resources which are internationally traded between soverign nation-states.

In a closed economy the optimal tax on an exhaustible resource is a pure rent tax, which is non-distortionary. We shall show that in some cases an ad valorem tax on output is equivalent to a rent tax. This raises the hope that a tariff on the import of an exhaustible resource may be optimal, but we shall show that in the conventional model of market equilibrium for an exhaustible resource the problem of finding the optimal tariff is not well-formulated, and raises fundamental and paradoxical questions. The problém arises because the importing country cannot enter into long term contracts with foreign suppliers, or more accurately, cannot be bound to such contracts, and can be expected to revise any tax plan it makes. Producers need to know the future tax intentions (or equivalently levels of import demand) in order to determine their current supply decision, but have no rational basis on which to predict the future. Optimal taxation or rational far-sighted producers is inconsistent.

5.1 The Taxation of Competitive Industry within National Boundaries

It is a familiar proposition of conventional tax theory that a tax on rent (or pure profits, correctly defined) is non-distortionary. This is also true for exhaustible resources, and the appropriate definition of rent turns out to be surprisingly simple.

Suppose the cost of extracting x units of exhaustible resources (called oil, for brevity) at date t, when the remaining stock is S is $C(x,S,t)$. Let the dollar value of consuming at rate x be $U(x,t)$ so the efficient consumption path solves

$$(1) \qquad \text{Max} \int_0^T [U(x,t) - C(x,S,t)]e^{-rt}dt + \phi(T)$$

subject to $x = -\dot{S}$, $\int_0^T x dt \leqq S_0$,

and given $C_{xx} \geqq 0$, $\partial C/\partial S \leqq 0$, $\partial C/\partial t \leqq 0$. ($T$ may be infinite, $\phi(T) = 0$, or T may be endogeneous, with $\phi(T)$ determining the present value of introducing a substitute at T.) Form the Hamiltonian

$$(2) \qquad He^{rt} = U - C - \mu x$$

where μe^{-rt} is adjoint to $\dot{S}$, and apply the Maximum principle to find the following necessary conditions:

$$\left.\begin{array}{r}\frac{\partial H}{\partial x} \leqq 0 \\ x \geqq 0\end{array}\right\}, \text{ or during the extraction phase } (x > 0)$$

(3a) $$p \equiv \frac{\partial U}{\partial x} = \frac{\partial C}{\partial x} + \mu = c + \mu$$

where c is the marginal extraction cost, $\partial C/\partial x$.

(3b) $$\dot{\mu} - r\mu = \frac{-\partial He^{rt}}{\partial S} = \frac{\partial C}{\partial S}$$

so

(4) $$\dot{p} = \frac{dc}{dt} + \frac{\partial C}{\partial S} + r(p - c)$$

With suitable concavity and boundary conditions this will uniquely determine the price trajectory, and it will be possible to decentralize production if resource owners are perfectly competitive and perfectly well informed about future prices. This can be seen as a special case of the impact of a rent tax (at zero rate) on such an industry.

If rents are defined as $px - C$, and taxed at a constant rate τ, producers will choose x to maximize

$$\int_0^T (1 - \tau)(px - C)e^{-rt}dt$$

subject to the same conditions as before.

(5) $$He^{rt} = (1 - \tau)(px - C) - \lambda x$$

(6a) $$p = c + \frac{\lambda}{1 - \tau}$$

and

(6b) $$\dot{\lambda} = r\lambda + (1 - \tau)\frac{\partial C}{\partial S}$$

Since if $\mu = \lambda/1 - \tau$ the equations are identical to (3), the same price equation results from eliminating λ, and, if the same boundary conditions are imposed, we have thus established

<u>Proposition 1</u>: Any constant <u>ad</u> <u>valorem</u> tax on rent defined as revenue less current extraction costs imposed on a competitive exhaustible resource industry leaves extraction Pareto efficient.

An excise tax on the output of a normal competitive industry is distortionary unless supply is completely inelastic, just as monopoly control of such an industry is distortionary. However, we know that under some conditions (constant elasticity of demand, zero extraction costs) a monopolized exhasutible resource may be efficiently extracted and the monopolist may have no monopoly power. It turns out that for essentially similar reasons, though under a wider range of conditions, an excise tax on oil may be nondistortionary. Indeed, in special cases an excise tax will be identical to a pure rent tax, allowing all the surplus to be taxed away.

Let p be the consumer (after tax) price, and p_n be the net producer price, so that the excise tax τ_e is $p - p_n$. Efficiency requires

$$(7) \qquad \dot{p} = \frac{\partial c}{\partial t} + \frac{\partial C}{\partial S} + r(p - c) \quad .$$

If competitive producers are to choose the efficient extraction path

$$(8) \qquad \dot{p}_n = \frac{\partial c}{\partial t} + \frac{\partial C}{\partial S} + r(p_n - c) \quad ,$$

both trajectories evaluated at the same values of $x(t)$. Subtracting we have

$$(9) \qquad \dot{\tau}_e = r\tau_e \quad \text{or} \quad \tau_e = \tau e^{-rt} \quad .$$

At each moment in time the present value of the stock of unexploited resource must be non-negative if producers are to continue extraction. Under the assumption made in equation (1) this is guaranteed if $p_n(T) \geq c(T)$, as the following Lemma shows.

<u>Lemma</u>: If $C_{xx} > 0$, $\partial C/\partial S \leqq 0$, then $p(t) \geqq c(t)$ implies $x(p(t') - c(t')) \geqq 0$, $t' < t$.

<u>Proof</u>: Suppose not, then $p(t') < c(t')$ and $x(t') > 0$. Reducing $x(t')$ to zero increases $S(t")$, $t" > t'$, reduces costs, increasing profits. Equivalently, note that $\mu \geqq 0$, $C_{xx} \geqq 0$ are sufficient to guarantee non-negative profits, and μ measures the value of having larger reserves of oil.

It follows that any tax on oil which satisfies equation (9) and the boundary condition $p_n(T) \geq c(T)$ is non-distortionary. Such a tax has a constant present value per unit of resource extracted, so the government is indifferent to the time path of extraction, and gains nothing from distorting this path. Unfortunately, the tax revenue may be negligible, or zero if the date at which the resource ceases to be extracted is determined by the marginal cost of extraction rising to the maximum demand price. If ever it is efficient to leave some resource unextracted, then the only non-distortionary excise tax is a zero rate tax; that is, excise taxes are then necessarily distortionary.

If $C_{xx} = 0$, or there are constant marginal extraction costs, then equation (4) simplified to

$$(10) \qquad \dot{p} = \frac{\partial c}{\partial t} + r(p - c)$$

and rents, $px - C$, become $(p - c)x$. We can now ask when an excise tax is equivalent to a rent tax, and equation (9) tells us that this can only be if $R = p - c$ rises at the rate of interest. From equation (4)

$$(11) \qquad \dot{R} = rR + \frac{\partial C}{\partial S}$$

which establishes

<u>Proposition 2</u>: If marginal extraction costs are independent of extraction rates, then an excise tax is equivalent to a rent tax (and thus non-distortionary) if extraction costs are also independent of resource stocks.

Intuitively it is easy to see why, for if $C_{xx} = \partial C/\partial S = 0$, the time path of extraction will not affect extraction costs at each date. Reducing all prices by the same present discounted amount does not provide any reason for reallocating resources over time on either the sales or cost side, and since the stock is a constant, overall supply is inelastic, and so, therefore, is supply at each date. Thus excise taxes are not distortionary, and the present value of the oil field can be almost entirely captured by the government.

5.2 Optimal Tariffs for Competitively Supplied Imports

The previous section analyzed taxes when the government had complete jurisdiction over producers and consumers. If oil is imported from foreign producers the government's tax powers are severely restricted to taxing imports, or, equivalently, restricting the level of imports. This limitation can profoundly alter the tax problem to the extent that there may be no optimal tax! The paradox arises in its most transparent form if extraction costs are independent of stocks and flows, and if every country has access to the same backstop technology which can provide unlimited supplies of energy at a constant cost $\bar{p}(t)$. Our country derives dollar benefits $U(x)$ from the consumption of x units of oil, and the demand by the rest of the world for oil of price p is $y(p)$. The problem is to choose a level of imports x, and production z from the backstop technology, to maximize

$$W = \int_0^\infty \{U(x + z) - px - \bar{p}z\}e^{-rt}dt \tag{12a}$$

subject to

(12b) $$-\dot{S} = x + y(p)$$

(12c) $$\dot{p} = \dot{c} + r(p - c) \quad .$$

Equations (12b) and (12c) describe the supply responses of the competitive producers who are assumed to perfectly forecast future demands and equilibrium prices. At some date T stocks of oil $S(T)$ will be exhausted and $p(T) = \bar{p}$, with the rest of the world switching to the backstop technology. The Hamiltonian is

(13) $$H = (U - px - \bar{p}z)e^{-rt} - \mu(x + y(p)) + \lambda(\dot{c} + r(p - c)) \quad .$$

Maximizing with respect to x, z:

(14a) $$\left.\begin{aligned} U' &\leqq p + \mu e^{rt} \\ x &\geqq 0 \end{aligned}\right\}$$

(14b) $$\left.\begin{aligned} U' &\leqq \bar{p} \\ z &\geqq 0 \end{aligned}\right\}$$

$$-\partial H/\partial S = \dot{\mu} = 0 \quad .$$

If q is the consumption price, $q = U'(x)$, then

(15) $$p = c + (p_0 - c_0)e^{rt}$$

$$q = \mathrm{Min}\ \{c + (q_0 - c_0)e^{rt}, \bar{p}\}$$

and the import tax $q - p$ is a rent tax, which has constant present value per unit of import.

Now for the paradox. If there were no other consumer, or, if all consumers formed a cartel and gave tax powers to a central agency, then the consuming countries could extract the entire surplus from the competitive producers (assuming they provoke no countervailing movement towards producer cartelization). If other consumers do not alter their demands $y(p)$ as our country varies its demand, then our country has to convince producers that it will set an initial excise tax μ, raise it at the rate of interest, and cease importing at the point where the consumer price reaches the backstop price $\bar{p}$. But the import price will still be below $\bar{p}$, and it will then be rational for the country to change its tax plan and continue to import until $p = \bar{p}$. In a world in which there are no binding contracts for future delivery there is no optimal tariff, because the consuming country will wish at each date to announce a time path of tax rates (or, equivalently, a time path of imports) which it will wish to change at the next date. The problem of choosing an optimal tariff is dynamically inconsistent in the Strotzian sense (Strotz [1955], so the producers have no rational basis on which to forecast future demands by the taxing consumer.

The result is so surprising and disturbing that it warrants closer examination. First, let us see quite clearly why, if producers believe what they are told about the tax intentions of the consuming country, the tax must rise at the rate of interest. Consider two adjacent time periods, $t = 1, 2$, and suppose that everyone is agreed on what is to happen from $t = 3$ onwards. Consider an increase in imports Δx at $t = 1$, and a reduction in imports Δx at $t = 2$, which leaves the future unchanged, and in particular, p_3 unchanged. It must also leave p_1 and p_2 unchanged, since they must satisfy the price equation (12c), or, more basically, because at the margin producers are indifferent to an intertemporal swap which leaves the future unchanged. The value to the consumer of the marginal change is

$$\Delta W = (q_1 - p_1)\Delta x - \frac{(q_2 - p_2)\Delta x}{1 + r}$$

and the consumer is optimizing if

$$\lim_{\Delta x \to 0} \frac{\Delta W}{\Delta x} = 0$$

or, $q - p$ rises at the rate of interest. But it will always be sensible for the consumer to import at any price below $\bar{p}$, _if_ he can do this without its having been foreseen. Producers will realize this, and not believe that taxes are going to always rise at the rate of interest, so they have no basis in which to make their price forecasts and hence current supply decisions. It will not even help for producers to assume that the optimal tariff is zero (the only tax rate which satisfies condition (15) _and_ allows imports until $p = \bar{p}$), for the consuming country can immediately invalidate this belief by imposing a current duty.

The structure of the paradox is akin to the paradox of the unexpected examination. The teacher announces that there will be an examination on one of the days from Monday to Saturday next week, and that it will be a surprise.

The students reason that this is impossible, for the exam cannot be a surprise on Staurday, so this day is ruled out, in which case the exam cannot be a surprise on Friday,.... The paradox is that the exam is set on Thursday, and is a surprise. In the tax paradox the producers reason that the final tax must be zero, and that it must rise at the rate on interest, in which case all previous taxes must be zero. But they won't be, because the country really does have some monopoly power.

The paradox seems to arise because we have assumed rational expectations on the part of producers and optimizing government. Is there some fundamental inconsistency in a non-stochastic world between rational expectations and optimal taxes?

5.3 Optimal Taxes and Rational Expectations

Suppose consumers live in a multi-period non-stochastic world in which it is costless to observe market behavior. Individuals are rational, and know the form of government's social welfare function, namely that it is some Bergsonian function of individual's revealed preferences. If goods and services are non-durable the optimum tax would seem to have the following structure. Optimize over market responses this period, and next period collect the same tax revenue from individuals as an individually specific lump sum tax. Modify this lump sum tax in the light of subsequent revealed market behavior. Of course, this tax will not be lump sum in effect, and might just as well have been a function of market behavior, since individuals will never believe that the apparently lump sum tax will not be modified. If the government could convince individuals in advance that it would not change its tax formulae (which on the above argument is reasonable) then it could collect information, which, if it could convince individuals it would now use to devise an unalterable individual lump sum tax, then everyone would be better off, but this subsequent promise is not believable. Some paradox is present, in that individuals will initially reason:

> "Suppose everyone deduces that the government is restricted to distortionary taxes which are a function of market behavior. Then everyone can optimize in the first period subject to these assumed unchanging tax functions. But in this case the government can do better than leave these rules unchanged, for it can set subsequent taxes as lump sum, and demonstrate that it is not prepared to change them. If this is ture, then I had better misrepresent my first period behavior to gain subsequent advantages under the lump-sum tax regime."

It would seem that we have the same paradox of the non-existence of optimal taxes. But here, the natural solution is that the government gives up any hope of imposing non-distortionary taxes, and there is an equilibrium in which consumers expect unchanging distortionary tax schedules, and learn nothing to cause them to modify this belief. That this situation should be described as a rational expectations optimal tax equilibrium seems reasonable, even if the temptation for the government to depart from it must be severe.

The paradox seems stronger when goods are durable, or when there are durable claims to future earnings like bonds. Suppose the rational expectations optimal tax involved no wealth taxation, then it would be optimal to impose an unexpected capital levy. If this is optimal, then rational consumers will forecast it, in which case it will not be unexpected (and thus non-distortionary), in which case it will not be optimal, so it will not be forecast, but then it will be unexpected, and hence optimal. The circularity seems vicious, and appears to have the same logical structure as the card on one side of which appears the sentence: "The statement on on the other side of this card is false," whilst on the reverse side is found, "The statement on the other side of this card is true."

The exhaustible resource is not only durable, but is locked into a highly structured temporal framework, for information about the past only affects the future up until the date when the stock is exhausted. The government can "punish" misrepresenting consumers by sticking to its original tax functions for all future time to ensure compliance, but the producers cannot indefinitely penalize misrepresenting consumers, as stocks run out in finite time.

5.4. Possible Resolutions of the Paradox

The problem of choosing an optimal tariff for oil imports is similar to that of choosing a strategy in an n-period Prisoner's Dilemma game. In this game each of two prisoners must decide between defection (D) and silence (S). If both choose S both receive mild sentences, if both defect both receive medium sentences, but if one defects and the other does not, he is freed and the other receives a particularly heavy sentence. The cooperative solution is silence, the dominant non-communicating strategy is to defect. In the n-period version it would seem profitable for each player to play S to indicate his willingness to implicitly collude but in period n each will fear that the other will renege, and will play D. It is then risky not to play D in n - 1, and so on, back to the present.

If, on the other hand, the game is played forever, or un unknown number of times, then it pays each player to implicitly collude and play S.

This suggests three possible resolutions of the tariff problem. If there is no backstop technology, then there is no dilemma point at which the importer will be tempted to abandon his tax plan. In general this is unlikely to avoid the problem of Strotzian inconsistency, for it seems likely that the "optimal" ad valorem tax rate calculated for any value of remaining reserves S_t will vary with S_t, but under the most favorable conditions it does seem possible to find a stationary solution with a constant ad valorem rent tax, as the following model suggests.

5.4.1 Model for Stationary Rent Tax

Assume zero extraction costs: $p(t) = pe^{rt}$, with $\bar{p} = pe^{rT}$ such that $S(T) = 0$.

Demand by the rest of the workd, $y(p) = p(t))^{-\varepsilon}$, and by the taxing country $x = a(p(t)\phi)^{-\varepsilon}$, where $U'(x) = p\phi$, $\phi = 1 + \text{tax rate}$. Suppose producers assume that ϕ will remain constant, and calculate the present price p given current reserves S from

$$S = \int_0^T (x + y)dt, \quad pe^{rT} = \bar{p}$$

$$S = p^{-\varepsilon}\int_0^T (a\phi^{-\varepsilon} + 1)e^{-\varepsilon rt}dt = \frac{(a\phi^{-\varepsilon} + 1)}{\varepsilon r}(p^{-\varepsilon} - \bar{p}^{-\varepsilon})$$

or

(16) $$p^{-\varepsilon} = \frac{\varepsilon r S}{a\phi^{-\varepsilon} + 1} + \bar{p}^{-\varepsilon} \quad .$$

Now suppose the taxing country deduces the form of equation (16) and chooses x to maximize instantaneous welfare subject to (16). The lagrangian is

$$L = U(x) - px + \lambda(\bar{p}^{-\varepsilon} - p^{-\varepsilon} + \frac{\varepsilon r S}{xp^{\varepsilon} + 1})$$

$$\frac{\partial L}{\partial x} \text{ gives } U' = p\phi = p + \frac{\lambda \varepsilon r S p^{\varepsilon}}{(a\phi^{-\varepsilon} + 1)^2}$$

$$\frac{\partial L}{\partial p} \text{ gives } \lambda = \frac{ap^{1-\varepsilon}\phi^{-\varepsilon}}{\varepsilon(p^{-\varepsilon} + \psi)}$$

where

(17) $$\psi = \frac{a\phi^{-\varepsilon}\varepsilon r S}{(1 + a\phi^{-\varepsilon})^2}$$

or

(18) $$\phi = 1 + \frac{\psi}{\varepsilon(\psi + p^{-\varepsilon})} \quad .$$

We now ask if ϕ could be constant, independent of S. For this

$$\psi = bS \quad , \quad b = \frac{a\phi^{-\varepsilon}\varepsilon r}{(1 + a\phi^{-\varepsilon})^2}$$

or

$$p^{-\varepsilon} = \gamma S \quad , \quad \gamma = \frac{b(1 + \varepsilon - \varepsilon\phi)}{\varepsilon(\phi - 1)} \quad .$$

But

$$p^{-\varepsilon} = \frac{\varepsilon r S}{(1 + a\phi^{-\varepsilon})} + \bar{p}^{-\varepsilon}$$

so in general (with $\bar{p}$ finite) it is inconsistent to assume ϕ independent of S. However, if $\bar{p} = \infty$, then a consistent solution is possible, provided (solving for γ)

$$\varepsilon(\phi - 1)\phi^{\varepsilon} = a$$

defines the optimal tax rate.

5.4.2 Uncertainty as a Resolution of the Paradox

The second possibility is that the stock of resource is unknown. Gilbert [1976] shows that if the probability distribution for remaining stocks is stationary, then the optimal extraction rate will also be constant (if costs and demands are also stationary). In this scenario a constant rent tax would be dynamically consistent, for the future will always look like the present. (To some extent this finesses the problem by making oil quasi-inexhaustible.)

The third possible resolution, and the one most in the spirit of conventional tax theory, is to suppose that the producers are uncertain of future demands, and make possibly probabilistic forecasts on the basis of past information, subsequently choosing a supply response to maximize expected present profits. If the taxing country knows the basis on which these forecasts are calculated it may then be able to choose an optimal import tariff.

The model of section 4.1 is in the spirit, and can be used to find the optimal tax as a function of remaining stocks:

$$\text{Max} \int \{U(x + z) - px - \bar{p}z\} e^{-rt}$$

subject to

$$S = (1 + xp^{\varepsilon}) \frac{(p^{-\varepsilon} - \bar{p}^{-\varepsilon})}{\varepsilon r} = \psi(x,p)$$

$$\dot{S} = -(x + y) \quad .$$

The Hamiltonian

$$H = U - px - p\bar{z} - \mu(x + y) + \lambda(S - \psi(x,p))$$

yields

$$U' \leqq p + \mu + (x + \mu y')(\frac{\psi_x}{-\psi_p})$$

and

$$\dot{\mu} - r\mu = \lambda = \frac{x + \mu y'}{\mu_p} \quad .$$

In principle these equations can be solved for x, though in practice it is difficult to find closed form solutions.

5.5 Conclusions

If exhaustible resources lie beyond the tax jurisdiction of a large consuming country then its attempt to tax imports optimally will have the following results when the resources are produced under competitive conditions. Taxes will be imposed at a positive rate, lowering consumption, leading to a more conservative use of resources, and conferrring external pecuniary benefits on other consumers at the expense of reduced producer rents. Not only will consumption be intertemporally inefficient but production will be inefficient, because rents will not be allowed to rise at the rate of interest. Producers' forecasts on which they base current supply decisions will be continually falsified by the tax policies of the importing country though both producers and the taxing country could be better off if they entered into binding long run supply contracts. In contrast, if there exists a supernational tax authority which can identify rents no distortion need arise from optimal taxation.

6. The Inefficiency Of the Competitive Stock Market and Its Implications For the Depletion Of Natural Resources

This appendix has two geneses. There has been some controversy associated with the optimality of the stock market allocation of resources. On the one hand, Diamond [1967] has argued that the stock market is a constrained Pareto Optimum, that it allocates resources as well as could a socialist economy in which the government were constrained to sharing the risks among the citizens by assigning to each a fraction of the profits of each firm (in addition to a fixed sum). It seems reasonable in evaluating the performance of a market economy in which there is an incomplete set of markets that one ought not to compare it (as say Borch has done) with how a socialist economy with a complete set of markets would have done; it is obvious that a market economy would not fare well in such an unfair comparison. Thus some constraint on the distribution of risks needs to be introduced, and that used by Diamond seems an appropriate one.

Ideally, one would like to have a theory which explains incompleteness of markets; several explanations have been offered associated with transactions costs and imperfections in information (see Stiglitz [1970], Grossman [1975]). Then an evaluation of the market economy would entail a comparison with a socialist economy facing the same transactions technology or the same information problems. In such a comparison, the market economy might be found deficient with respect to the set of markets which operate. The discussion of this paper can be viewed as a "minimal" test of the market economy: given the set of markets operating does it allocate resources efficiently; if it does, it still does not mean that one should view the market economy as being efficient, but if it does not, it clearly can be viewed as being inefficient.

Stiglitz [1972](see also Jenson-Long [1972]) on the other hand, has argued that the stock market is not a constrained Pareto Optimum. Much of the discussion has revolved around what are the appropriate behavioral assumptions for the firm; do firms value maximize and do they believe that if they double their scale they double their market value. (See Stiglitz [1970b], Grossman-Stiglitz [1976]). For the Diamond results to obtain, there must be a large number of firms with perfectly correlated returns; and the firm must have what Diamond calls, multiplicative uncertainty, i.e., doubling inputs increases output in each state of nature in proportion. Under these assumptions, all stockholders will wish the firm to maximize market value [A.3/] and bankruptcy [A.4/] (which in general plays an important role) can be ignored.

What I wish to show here is that even under these seemingly favorable conditions for the market economy, the stock market allocation of resources is not a constrained Pareto Optimum when there are more than two outputs. The reason for this is that the constraint on the distribution of profits involves prices; although in a competitive economy each firm will ignore its effects on prices, a socialist economy would not.

The matter may be put another way. With a single output there is no distinction between stochastic homotheticity for a firm or for an industry: the ratio of output or profits in any two states of nature remains a constant. With more than one commodity, if a single firm increases its scale, it may perceive that the ratio of its profits in any two states of nature remains constant; but when the industry expands its scale, the price of the output of the industry may change differently in different states, so that the ratio of industry profits in two different states changes as the scale of the industry changes; there is private stochastic homotheticity -- taking prices in each state of nature as given -- but not social stochastic homotheticity. The former assures that firms will wish to maximize stock market value and makes it reasonable for them to assume that value is proportional to their scale. But the latter is what is ncesssary to ensure constrained optimality of the market solution.

Although Stiglitz [1973] and Hart had earlier noted the non-optimality of the stock market allocation when there were more than one output, they provided a less complete characterization of the source of the inefficiency than we are able to provide here.[A.5/]

This brings me to the second motivation for this paper.

The decision to consume oil (or any other natural resource) today or to postpone consumption until next period is an investment decision. It is a risky investment decision, since the return to holding the oil is random; it depends on the price next period, which in turn depends on the amount of oil that will be discovered. It has been argued that because of the absence of a complete set of risk markets, the economy behaves in a more risk averse manner than is socially optimal; and because holding oil is risky, this implies that too little oil will be held over, i.e.

that the market will result in too fast depletion of our resources. This argument is, however, unconvincing on two grounds. First, if there are incomplete risk markets, the risks which are borne by holders of oil -- which in a complete set of markets they might be able to insure against -- have a real social cost; that is, the fact that the risks might, under some other institutional arrangement, be spread does not alter the fact that those who hold oil, in our present institutional arrangements, are bearing a risk which affects their expected utility. The conventional discussions of the issue seem to make the error of comparing the allocation of resources with one market structure with the allocation which would have emerged with another market structure, a comparison which we have argued earlier is both unfair, and probably irrelevant.

The conventional argument is unpersuasive on a second account: there is a distinction between consumer risk and producer (or investment) risk. Because of the variability in the price of oil a consumer would face risk even were his income (denominated in either oil or "other goods") constant. Only if an individual were perfectly hedged, i.e. owned, in each state of nature, exactly the amount of oil he consumed, would the price variability be of no concern. In general, both because of differences in tastes and in risk aversions, individuals will not be perfectly hedged (either in the market equilibrium or in the optimal allocation of resources). When individuals are not perfectly hedged, we might think of there being two groups, those who own oil (or shares in oil firms) and those who consume oil. If the former are, say, less risk averse than the latter, then holding oil is privately more attractive than it is socially desirable, and there is excessive conservation.

The Basic Model

There are two periods and two commodities. The two commodities are oil and "other consumption goods." The consumption of "other consumption goods" the first period we denote by C_1, the second period by C_2, and we assume that there is a fixed stock of C which can be consumed in one or the other period:

(1) $$C_1 + C_2 = \bar{C} \quad .$$

(In effect, the marginal rate of transformation of C between this period and next is assumed to be unity.)

Consumption of oil the first period is denoted by Q_1 and is constrained by the stock of known reserves

(2) $$Q_1 \leq \bar{Q}_1 \quad .$$

Next period we can consume that part of known reserves which we have not consumed this period; in addition, it is known that under certain tracts of land there is oil, but the amount there is unknown. Whatever is there will, however, be consumed next period. Thus

(3) $$\tilde{Q}_2 = \bar{Q}_1 - Q_1 + \tilde{Q}_d \equiv \tilde{Q} - Q_1$$

where the $\tilde{}$ above $\tilde{Q}_d$ and $\tilde{Q}$ is to remind us that the variable in question is random (at the time at which the decisions are being made) and where $\tilde{Q}$ is total oil stocks and Q_d is oil discoveries,

$$\tilde{Q} = \bar{Q}_1 + Q_d$$

and $\tilde{Q}_2$ is consumption of oil next period.

There are three assets in the economy: known oil stocks $(\bar{Q}_1)$, "other" consumption goods, C, and claims on future discoveries of oil. Let $\{\alpha_0^j, \beta_0^j, \gamma_0^j\}$ represent the shares of each of these stocks owned initially by the jth individual. Thus, we let C be our numeraire, p_1 be the price of oil (the first period), and s_1 be the total market value today of the unknown future reserves, then the jth individuals' initial wealth, W^j, is given by

$$(4) \qquad W^j = \alpha_0^j C + \beta_0^j \bar{Q}_1 p_1 + \gamma_0^j s_1 \quad .$$

The price of oil next period is a random variable, denoted by $\tilde{p}$. The individual has expectations about $\tilde{p}$ as well as about $\tilde{Q}$. He allocates his initial wealth about consumption of oil this period, consumption of other goods this period, ownership claims in the safe asset (C), ownership claims in known oil stocks, and ownership claims on unknown oil reserves, in order to maximize his expected utility. For simplicity, we assume his utility function is separable:

$$(5) \qquad U^j = U_1^j(C_1^j, Q_1^j) + \frac{1}{1+\delta} EU_2^j(C_2^j, Q_2^j)$$

$$= U_1^j(C_1^j, Q_1^j) + \frac{1}{1+\delta} EV^j(\tilde{p}, \tilde{Y}^j)$$

where V^j is the indirect utility function corresponding to U_2^j, and Y^j is second period wealth. If Z^j is the jth individuals holding of a safe asset (with the interest rate equal to r), β^j is his ownership share is known oil stocks which are not consumed first period, and γ^j is his ownership share in unknown oil reserves,

(6) $$\tilde{Y}^j = Z^j(1 + r) + \beta^j\tilde{p}(\bar{Q}_1 - Q_1) + \gamma_1^j\tilde{Q}_d\tilde{p} \quad .$$

The jth consumer

(7) $$\underset{\{C_1^j,\ Q_1^j,\ Z^j,\ \beta^j,\ \gamma^j\}}{\text{maximizes}} \quad U^j$$

s.t.

$$W^j \geq C_1^j + p_1Q_1^j + Z^j + \beta^j p_1(\bar{Q}_1 - Q_1) + \gamma^j s_1 \quad .$$

We can immediately write down the first order conditions:

(8a) $$\frac{\partial U_1^j}{\partial C_1^j} = \frac{(1 + r)}{1 + \delta} EV_Y^j = \frac{(1 + r)}{1 + \delta} E \frac{\partial U_2^j}{\partial C_2^j} \quad ,$$

the generalization to uncertainty of the familiar condition that the marginal rate of substitution between consumption this period of "C" and consumption next period of "C" is equal to one plus the rate of interest;

(8b) $$\frac{\partial U_1^j}{\partial C_1^j} = \frac{\partial U_1^j}{\partial Q_1^j}/p_1$$

the marginal rate of substitution between consumption of oil and of C equals the price ratio;

(8c) $$\frac{\partial U_1^j}{\partial Q_1^j} = \frac{1}{1+\delta} EV_Y^j \tilde{p} \quad ;$$

substituting (8a) and (8b) into (8c), we obtain

(8c') $$p_1(1 + r) = \frac{EV_Y^j \tilde{p}}{EV_Y^j} \quad .$$

This is the generalization to the uncertain context of the familiar condition that the price of a natural resource must rise at the rate of interest. Here, we have the next period's weighted expected price, where the weights are marginal utilities of income, is equal to today's price times $1 + r$.

Finally, we have the portfolio equilibrium condition

(8d) $$\frac{EV_Y^j \tilde{p}\tilde{Q}_d}{EV_Y^j} = s_1(1 + r) \quad .$$

Note that from (8c') and (8d), the weighted average value of next period's price must be the same for everyone, as well as the weighted average value of the discoveries next period:

(8e) $$\frac{EV_Y^j \tilde{p}}{EV_Y^j} = p_1(1 + r) \quad \text{for all } j$$

(8f) $$\frac{EV_Y^j \tilde{p}\tilde{Q}_d}{EV_Y^j} = s_1(1 + r) \quad \text{for all } j$$

The solution to (8) yields demand functions of the form

(9) $$C^j = C^j(p_1, s_1, 1 + r, \tilde{p}; \alpha_0^j, \beta_0^j, \gamma_0^j)$$

$$Q_1^j = C^j(p_1, s_1, 1 + r, \tilde{p}; \alpha_0^j, \beta_0^j, \gamma_0^j)$$

$$Z^j = Z^j(p_1, s_1, 1 + r, \tilde{p}; \alpha_0^j, \beta_0^j, \gamma_0^j)$$

$$\beta^j = \beta^j(p_1, s_1, 1 + r, p; \alpha_0^j, \beta_0^j, \gamma_0^j)$$

$$\gamma^j = \gamma^j(p_1, s_1, 1 + r, p; \alpha_0^j, \beta_0^j, \gamma_0^j)$$

Market equilibrium requires

(10) $$\Sigma(C^j + (1 + r)Z^j) = C$$

$$\Sigma Q_1^j = Q_1 \leq \bar{Q}_1$$

$$\Sigma\beta^j = 1$$

$$\Sigma\gamma^j = 1$$

So far, we have said nothing about how expectations about next periods prices are formed. Obviously, if individuals expectations are incorrect, there will be an "incorrect" allocation of resources. We ask, suppose individuals could perfectly predict the probability distribution of prices, i.e. they had rational expectations, would the market yield a constrained Pareto Optimum. Surely, if the market does not work well with rational expectations, there is no reason to believe that it will work well with arbitrarily formed expectations. In this sense, imposing the constraint of rational expectations biases the case in favor of the market.

In this context, rational expectations requires that the price next period, conditional on the realization of Q_d, be equal to the actual price, i.e.

(11) $$E\{\tilde{p}|Q_d\} = p(Q_d),\ \text{Var}\ \{\tilde{p}|Q_d\} = 0$$

where $p(Q_d)$ is derived as follows: given the portfolio allocation decision of the individual, for any realization of Q_d, we have

(12) $$Y^j(Q^d) = Z^j(1 + r) + \beta^j p(Q_d)(Q_1 - \bar{Q}_1) + \gamma^j p(Q_d)Q_d \quad .$$

The consumers demand for oil is given by

(13) $$Q_2^j \equiv - \frac{V_P^j(p, Y^j)}{V_Y^j(p, Y^j)} \quad .$$

$P(Q_d)$ must be such that

(14) $$\Sigma Q_2^j = \Sigma \frac{V_P^j}{V_Y^j} = Q_1 - \bar{Q}_1 + Q_d$$

This model can be viewed in several alternative ways. We can pretend that there are firms which hold present oil stocks. A firm which "invests" in oil by holding an amount $\hat{Q}$ will have profits to distribute next period $\tilde{p}\hat{Q}$. The cost of this investment is $p\hat{Q}$. Thus the firm does face stochastic constant returns to scale: if it doubles the amount it invests it doubles its profits in every state of nature. Value maximizing firms will continue to expand (invest in oil) to the point where the value of the firm is exactly equal to the value of its investment; for if the value of the firm exceeds the value of its investment, it believes by doubling its investment, it doubles its value, and the "net value" (the value of the firm minus its investment, which is equal to the value of the equity of the original shareholders) will be increased. But as all firms do this, the price next period (in each state of nature) will go down, and the price of oil this period will increase, until equilibrium is attained.

Since all firms holding oil are perfect substitutes, we do not need to distinguish among them, and β^j represents the jth individuals share in the entire industry. Similar remarks apply to firms which invest in tracts of land under which oil is located.

It is clear in this persepctive, the model is a direct generalization of the Diamond model to many commodities, but we have simplified the technology so that all firms face (what they view as) stochastic constant returns (rather than the multiplicative uncertainty used by Diamond). The constraint on the distribution of profits that we have imposed is identical to the constraints imposed in his model, except that profits in his model were technologically determined. In our model, profits of any holder of oil stocks depend on the price, which depends on the discoveries of oil (a technological parameter for the exploration industry); and for the firm in exploration industry, profits depend not only on the amount of oil discovered by a particular firm, but also on price; but price depends on the amount of oil discovered by all firms.

Alternatively, this model can be viewed as a pure exchange model, with futures markets for the price and the value of future discoveries of oil.

We now contrast the competitive rational expectations equilibrium with a constrained Pareto Optimum, which can be viewed as the solution to the following maximization problem.

$$\max_{\{C_1^j, Q_1^j, X_1^j, \beta^j, \gamma^j, Q_1\}} \Sigma\lambda^j U^j = \Sigma\lambda^j U_1^j(C_1^j, Q_1^j) + \frac{\lambda^j}{1+\delta} EV^j(p, X^j + \beta_j(\bar{Q}_1 - Q_1)\tilde{p} + \gamma^j \tilde{Q}_d \tilde{p}) \tag{15}$$

s.t.

$$\Sigma C_1^j + \Sigma X^j = C \tag{16}$$

$$\Sigma\beta_j = 1$$

$$\Sigma\gamma_j = 1$$

$$\Sigma Q_1^j = Q_1 \leq \bar{Q}_1$$

and where, as before, we assume that price expectations are rational, i.e. satisfy (11 - 14), where now

$$Y^j = X^j + \beta^j(\bar{Q}_1 - Q_1)\tilde{p} + \gamma^j\tilde{Q}_d\tilde{p} \tag{17}$$

We assume, in other words, that the government can allocate purchasing power the first period (C_1^j), oil for consumption the first period (this is a redundant instrument, if we allow a competitive market the first period and trading of C for oil); but that the government is constrained in the instruments by which it can spread risk in the same manner that the market is, i.e. it can issue non-negotiable shares in the unknown oil reserves, it can open up a futures market for oil (since $\bar{Q}_1 - Q_1$ is a known constant, the β^j represent simply a gamble on the future price of oil); and it can distribute a safe bond (X^j).

The first order conditions may be easily written down:

$$\frac{\lambda^j \partial U_1^j}{\partial C_1} = \frac{\lambda^k \partial U_1^k}{\partial C_1} \tag{18a}$$

$$\frac{\lambda^j \partial U_1^j}{\partial Q_1} = \frac{\lambda^k \partial U_1^k}{\partial Q_1} \tag{18b}$$

From (18a) and (18b) we immediately obtain

$$\frac{-\frac{\partial U_1^j}{\partial C_1}}{\frac{\partial U_1^j}{\partial Q_1}} = \frac{\frac{\partial U_1^k}{\partial C_1}}{\frac{\partial U_1^k}{\partial Q_1}} \tag{18c}$$

the marginal rate of substitutes between oil and C this period must be the same for all individuals; this corresponds to condition (8b) for the competitive economy.

(18d) $$\lambda^j EV^j_{2Y} = \lambda^k EV^k_{2Y}$$

From (18a) and (18d), we obtain

$$\frac{EV^j_{2Y}}{\frac{\partial U^j_1}{\partial C_1}} = (1+\delta)\frac{E\frac{\partial U^j_2}{\partial C_2}}{\frac{\partial U^j_1}{\partial C_1}} = (1+\delta)\frac{E\frac{\partial U^k_2}{\partial C_2}}{\frac{\partial U^j_1}{\partial C_1}} = \frac{EV^k_{2Y}}{\frac{\partial U^k_1}{\partial C_1}}$$

which corresponds to the condition (8a) for the competitive economy;

(18e) $$\lambda^j EV^j_{2Y}\tilde{p} = \lambda^k EV^k_{2Y}\tilde{p}$$

which, with (18d) yields

$$\frac{EV^j_Y\tilde{p}}{EV^j_Y} = \frac{EV^k_Y\tilde{p}}{EV^k_Y}$$

corresponding to (8c') of the competitive economy.

Similarly, from the first order condition for γ^i_j we obtain

(19g) $$\lambda^j EV^j_{2Y}\tilde{p}\tilde{Q}_d = \lambda^k EV^j_{2Y}\tilde{p}\tilde{Q}_d$$

implying that

(18h) $$\frac{EV^j_{2Y}\tilde{p}\tilde{Q}}{EV^j_{2Y}} = \frac{EV^k_{2Y}\tilde{p}\tilde{Q}}{EV^k_{2Y}}$$

corresponding to (8d) for the competitive economy.

So far, all the conditions for constrained pareto optimality have their perfect analogue in the market economy. But the final condition, relating to total savings of oil for next period $(\bar{Q}_1 - Q_1)$ is different; for by changing Q_1, aggregate consumption of oil first period, we change the supply of oil next period, which has an important effect on the price distribution next period, p; this affect on the the price distribution is what each competitor ignores, but which the government, in solving the constrained pareto optimality problem, does not. We thus have , choosing Q_1)

(18i) $$\frac{1}{1+\delta}\sum_j \lambda^j\{E[[V^j_p + V^j_Y \beta_j(\bar{Q}_1 - Q_1) + \gamma_j\tilde{Q}_d]\frac{dp}{dQ_1}] + V^j_Y \beta_j\tilde{p}\} - \frac{\partial U^0_1}{\partial Q_1}\lambda^0 = 0$$

From (18e) we obtain

$$\Sigma\lambda^j_E V^j_{2Y}\tilde{p}\beta^j = \lambda^0 EV^0_{2Y}\tilde{p}\Sigma\beta^j$$

$$= \lambda^0 EV^0_{2Y}\tilde{p} \quad .$$

In competitive equilibrium (equation 8c)

$$\frac{1}{1+\delta} EV^j_{2Y}p = \frac{\partial U^j_1}{\partial Q_1} \quad .$$

Hence, the market economy is a constraincd pareto optimum if and only if (using the well known result that $Q_2^j = -V_p/V_Y$)

$$E \frac{d\tilde{p}}{dQ} \lambda^j [V_p^j + V_Y^j \beta_j (\bar{Q}_1 - Q_1) + \gamma_j Q_d] = -E \frac{d\tilde{p}}{dQ} \Sigma \lambda^j V_Y^j \Delta^j = 0$$

where

(20) $$\Delta^j \equiv Q_2^j - \beta^j (\bar{Q}_1 - Q_1) - \gamma^j \tilde{Q}_d \quad ,$$

the difference between the jth individuals second period consumption of oil and his (implicit) second period endowment of oil. (If the individual actually stored his own oil and bought his own tract of land with unknown oil reserves, then $\beta^j (Q_1 - \bar{Q}_1) + \gamma^j \tilde{Q}_d$ would be his actual holdings of oil; otherwise we can think of oil firms as distributing profits by delivering their oil stocks; the individual could then trade these stocks. Of course it makes no difference whether the individual trades the oil for dollars himself or the firm does it for him.)

Δ^j is, in other words, his net trade in oil second period. If the individual were perfectly hedged,

$$\Delta^j = 0$$

i.e. his purchases this period of oil for consumption next period are exactly equal to his consumption. Ex post, it has turned out that he was neither speculating on an increase in the price of oil (he purchased more than he consumes, $\Delta^j < 0$) or speculating on a fall in the price of oil (he purchased less than he consumes, $\Delta^j > 0$).

We immediately obtain the result that a sufficient condition for the constrained pareto optimality of the market is that everyone is perfectly hedged ($\Delta^j \equiv 0$ all j).

From (17), it is clear that

$$(21) \qquad \Sigma\Delta^j = \Sigma Q_2^j - \Sigma\beta^j(\bar{Q}_1 - Q_1) - \Sigma\gamma^j Q_d = Q_2 - (\bar{Q}_1 - Q_1) - Q_d = 0$$

total aggregate consumption second period must equal the supply of oil; hence, if all individuals are identical, $\Delta^j \equiv 0$, and the market is a constrained Pareto optimum.

An alternative sufficient condition for the constrained Pareto optimality of the stock market is that, for each state of nature (realization of Q_d),

$$(22) \qquad \lambda^j V_Y^j = \lambda^k V_Y^k$$

(Constrained Pareto Optimlaity assures us only that

$$E\lambda^j V_Y^j = E\lambda^k V_Y^k \quad .)$$

In other words, if we could make state dependent transfers of income, we would not wish to do so, then the market, again not surprisingly, yields a constrained Pareto Optimum. An obvious case of (22) arises with constant marginal utility of income with the constant being inversely proportional to λ^k. This, however, is not a very interesting case, for if all individuals consume both commodities, there cannot be constant marginal utility of income unless, by chance, the equilibrium involves no price variability:

$$\frac{dV^j_{2Y}}{dp} = \frac{\partial V^j_{2Y}}{\partial p} + \frac{\partial V^j_{2Y}}{\partial Y^j}\frac{\partial Y^j}{\partial p}$$

$$= -V^j_Y \frac{\partial Q^j_2}{\partial Y} - Q^j_2 V^j_{2YY} + V^j_{2YY}(\beta^j(\bar{Q}_1 - Q_1) + \gamma^j Q_d)$$

$$= \frac{Q^j_2 V^j_Y}{Y^j}\left[\frac{R^j_2 \Delta^j}{Q^j_2} - \eta^j\right]$$

where

$$R^j_2 = -\frac{V_{2YY}Y^j}{V^j_Y} \quad \text{, the measure of relative risk aversion}$$

$$\eta^j = \frac{\partial \ln Q^j_2}{\partial \ln Y^j} \quad \text{, the income elasticity of demand for oil}$$

Since if $\Sigma \Delta^j \equiv 0$, for some j, $\Delta^j < 0$, so $dV^j_Y/dp <$ if $\eta^j > 0$.

In general, however, (19) is not satisfied, and thus the <u>market equilibrium is not a constrained Pareto Optimum</u>. A further characterization of the nature of the biases in the market allocation requires some simplification of the model. In the subsequent sections of the paper, we consider some special cases in which the nature of the bias may be more precisely assessed.

Approximation Formulae for Market Bias

We assume, in addition, that holdings of known oil stocks are in proportion to holdings of undiscovered oil, i.e. in our notation

$$\beta^j = \gamma^j \quad . \tag{24}$$

It can be shown that for small variances, whenever condition (8e) is satisfied, condition (8f) is also satisfied (to a second order of approximation); thus, for symmetric distributions, differences between β^j and γ^j are a function of fourth order and higher terms; thus for practical purposes we can impose the restriction $\beta = \gamma$. (Alternatively, under certain simple structures pQ_d is a linear function of p and Q_d, and the market for reserves is redundant.)[A.6]

For simplicity, we assume only two groups, denoted by superscripts a and b. Then, for small variations, for (8e) to be satisfied[A.7]

$$\frac{V^a_{YQ}}{V^a_Y} = \frac{V^b_{YQ}}{V^b_Y} \quad . \tag{39}$$

In addition, we will assume optimal lump sum redistributions, so

$$\lambda EV^a_Y = (1 - \lambda)EV^b_Y$$

or again, for small variations,

$$\lambda V^a_Y(\bar{Q}) - (1 - \lambda)V^b_Y(\bar{Q}) = [(1 - \lambda)V^b_{YQQ} - \lambda V^a_{YQQ}]\frac{\sigma^2_Q}{2} \quad . \tag{39'}$$

Thus (19) can be approximated by

(40) $$Ep'\{V_Y^a\lambda - (1-\lambda)V_Y^b\}\Delta \approx \Delta(\bar{Q})p'(\bar{Q})[(1-\lambda)V_{YQQ}^b - \lambda V_{YQQ}^a]\frac{\sigma_Q^2}{2} .$$

We can rewrite (40) as

$$D = (\frac{\Delta}{Q})(\frac{p'Q}{p})p[(1-\lambda)V_{YQQ}^b - V_{YQQ}^a]\sigma_Q^2 .$$

The magnitude of the distortion in Q can thus be approximated by letting $W(Q_2)$ be the value of social welfare as a function of $\bar{Q}_2$:

(41) $$\frac{\bar{Q}_2^0 - \bar{Q}_2^m}{Q_2^0} = - \frac{D}{W''Q_2^0}$$

where $\bar{Q}_2^0$ is the optimal value of $\bar{Q}_2$ and $\bar{Q}_2^m$ is the value in the market solution.

Essentially (40) measures the magnitude of the income effect (transfer) from the price change induced by additional savings of oil. Normally, $p' < 0$, i.e. an increase in the quantity of oil reduces its price. Ascertaining the sign of Δ and $(1-\lambda)V_{YQQ}^b - \lambda V_{QQ}^a$ is more difficult. We consider here only the case of $\lambda = (1-\lambda)$.

It is convenient at this point to focus on the special case of homothetic utility functions. Then

$$\frac{\partial Q_c^a}{\partial Y^a} = \frac{Q_c^a}{Y^a} \equiv \frac{S_c^a}{p} ,$$

when

$$S_c^a = \text{share of a's expenditure on consumption of oil (second period)}$$

Then, using (31) and (33) we calculate

(42) $$2\,\frac{\Delta}{Q}vR \equiv \frac{v\hat{S}_c}{\bar{S}} - \hat{R}$$

where

$$S_c \equiv S_c^a - S_c^b$$

$$\hat{R} \equiv \frac{\gamma R^a}{\zeta} - \frac{(1-\gamma)R^b}{1-\zeta}$$

$$2R = \frac{R^a}{\zeta} + \frac{R^b}{1-\zeta}$$

Since $Q_c^a(p,Y^a) + Q_c^b(p,Y^b) = Q_2$

when Q_c^j is the j^{th} individual's demand function for oil second period,

(43) $$\frac{p'Q}{p}[-\bar{\varepsilon} - \frac{\Delta}{Q}\hat{S}_c] + \bar{S}_c = 1$$

where

$\bar{S}_c = \gamma S_c^a + (1-\gamma)S_c^b$, a weighted average of consumption shares in oil

$\varepsilon = -(\frac{\partial \ln Q}{\partial \ln p})_{\bar{u}}$, the compensated price elasticity

$\bar{\varepsilon} = \frac{Q_c^a}{Q_2}\varepsilon^a + \frac{Q_c^b}{Q_2}\varepsilon^b$, a weighted average compensated price elasticity

We rewrite (42) and (43) as

(43') $$v\bar{\varepsilon} + \frac{v\Delta}{Q}\hat{S}_c = 1 - \bar{S}_c$$

(42') $$-v\frac{\hat{S}_c}{\bar{S}} + 2v\frac{\Delta}{Q}R = -\hat{R} \quad .$$

Hence

(44) $$\frac{v\Delta}{Q} = \frac{\frac{(1 - \bar{S}_c)}{\bar{\varepsilon}}\frac{\hat{S}_c}{\bar{S}} - \hat{R}}{2R + \frac{\hat{S}_c}{\varepsilon}\frac{\hat{S}_c}{\bar{S}}}$$

(45) $$v = \frac{2[(1 - \bar{S}_c)R + \hat{R}\hat{S}_c]/\bar{\varepsilon}}{2R + \frac{\hat{S}_c^2}{\bar{S}\bar{\varepsilon}}}$$

(46) $$\frac{\Delta}{Q} = \frac{(1 - \bar{S}_c)\frac{\hat{S}_c}{\bar{S}} - \hat{R}\bar{\varepsilon}}{2(1 - S_c)R + \hat{R}S_c}$$

If individuals are similar, so $\hat{S}_c$ and $\hat{R}$ are small, then

(47) $$v \approx \frac{1 - \bar{S}_c}{\bar{\varepsilon}}$$

(48) $$\frac{\Delta}{Q} \approx \frac{\hat{S}_c}{2R\bar{S}} - \frac{\hat{R}\bar{\varepsilon}}{2R(1 - \bar{S}_c)}$$

If risk aversions are identical and if we have normalized our utility functions so with optimal lump sum transfers, $Y^a(\bar{Q}) \approx Y^b(\bar{Q})$, then (46) can be rewritten

$$(1 - \bar{S}_c)\frac{\Delta}{Q} \times 2\bar{S}R \approx (1 - \bar{S}_c)\hat{S}_c + 2R\bar{\varepsilon}(1 - 2\gamma)\bar{S} \quad .$$

But recalling that

$$\Delta = Q_c^a - \gamma Q$$

$$\gamma = \frac{S_c Y^a}{PQ} - \frac{\Delta}{Q}$$

$$1 - 2\gamma = 1 + \frac{2\Delta}{Q} - \frac{2Q_c^a}{Q} = \frac{Q_c^b - Q_c^a}{Q} + \frac{2\Delta}{Q} \quad .$$

Thus

(49) $$\frac{\Delta}{Q} \approx \frac{[1 - \bar{S}_c - \bar{\varepsilon}R]\hat{S}_c}{2R\bar{S}(1 - \bar{S}_c - 2\bar{\varepsilon})} \quad .$$

The individual who consumes more (relatively) oil acts as if he were more risk averse; i.e. the other individual speculates on the oil.

We now need to ascertain the sign of

$$\frac{dV^a_{YQ} - V^b_{YQ}}{dQ} = \frac{V^a_{YQ}}{V^Q_Y}\frac{dV^a_Y - V^b_Y}{dQ}$$

$$+ V^A_Y \frac{dV^a_{YQ}/V_Y}{dQ} - V^B_Y \frac{dV^b_{YQ}/V_Y}{dQ}$$

$$\approx - \frac{p}{Y} V_y \left\{\frac{d\{2\frac{\Delta}{Q} vR + \hat{R} - \frac{v\hat{S}_c}{\bar{S}}\}}{dQ}\right\} \quad .$$

We observe that

(51)
$$\frac{Qd\Delta/Q}{dQ} = \frac{\varepsilon^a Q^a_c v}{Q} + \frac{\Delta}{Q}(S^a_c v - 1) - \gamma(1 - S^a_c)$$

$$= \frac{Q^a_c}{Q}(1 - \bar{S}_c) - \gamma(1 - S^a_c) + \frac{\Delta}{Q}(S^a_c v - 1) + \frac{Q^a_c}{Q}\left(\frac{Q^b_c}{Q}\right)(\varepsilon^b - \varepsilon^a)v$$

$$= S^a_c(v - 1)\frac{\Delta}{Q} + (1 - \gamma)\hat{S}_c \frac{Q^a_c}{Q} + \frac{Q^a_c}{Q}\left(\frac{Q^c_c}{Q}\right)(\varepsilon^b - \varepsilon^a)v$$

(52)
$$\frac{dv}{dQ} = \left[\frac{vd\bar{\varepsilon}}{dQ} + vS_c\frac{d\Delta/Q}{dQ} + v\frac{\Delta}{Q}\frac{d\hat{S}_c}{dQ} + \frac{d\bar{S}_c}{dQ}\right]/\bar{\varepsilon} + \frac{\Delta}{Q}\hat{S}_c$$

(53)
$$\frac{d\hat{R}}{dQ} \approx -\zeta'\left[\frac{\gamma R^a}{\zeta^2} + \frac{(1 - \gamma)R^b}{(1 - \zeta)^2}\right]$$

where

$$\zeta' = \left(\frac{\gamma Y - Y^a}{Y^2}\right)(1 - v)p = \left(\frac{\gamma - \zeta}{Y}\right)(1 - v)p$$

$$= \left(\frac{\hat{S}_c \zeta(1 - \zeta)}{\bar{S}} - \frac{\Delta}{Q}\right)(1 - v)\frac{p}{Y} \quad .$$

$$(54) \qquad Q\frac{d\hat{S}_c}{dQ} = -v(\hat{S}_c(1-(S^a+S^b)) - (S^a\varepsilon^a - S^b\varepsilon^b))$$

$$(55) \qquad \frac{d\bar{S}_c}{dQ} = -v\bar{S}\left(1 - \frac{\gamma S^{a2} + (1-\gamma)S^{b2}}{\bar{S}} - \frac{\varepsilon^a\gamma S^a + \varepsilon^b(1-\gamma)S^b}{\bar{S}}\right)$$

For similar individuals (small Δ/Q) with first and second period utility functions being approximately the same, we obtain

$$\frac{Q}{V_Y}\frac{\partial^2 W}{\partial Q_1^2} \approx -\frac{\partial U_1/\partial Q_1^2 + \partial^2 U_2/\partial Q_2^2}{V_Y}Q$$

$$\approx \frac{2p[(1-\bar{S}_c) + \bar{R}\bar{S}_c]}{\varepsilon - S_c}$$

(We make use of the fact that

$$U_Q = pV_Y$$

so

$$U_{QQ}dQ = dp[V_Y + pV_{Yp}]$$

$$= dp[V_Y(1-S_c) + V_Y RS_c]$$

$$QU_{QQ} = V_Y p\frac{(1-S_c) + RS_c}{\varepsilon - S_c} \quad .)$$

Hence

$$(56) \qquad \frac{\bar{Q}_2^0 - \bar{Q}_2^m}{\bar{Q}_2^0} \approx \frac{\frac{\Delta}{Q}\; v \frac{d\left(\frac{V_{YQ}^a}{V_Y} - \frac{V_{YQ}^b}{V_Y}\right)\sigma_Q^2(\varepsilon - \bar{S}_c)}{dQ}}{2[(1 - S_c) + RS_c]Q_2}$$

To ascertain the sign and magnitude of (55) we simply substitute from (45), (46), (50), (51), (52) into (53). The expressions obtain appear messy and not easy to interpret.

The analysis does, however, identify all the important parameters determining both the magnitude and the direction of the bias in the intertemporal allocation of oil.

APPENDIX

FOOTNOTES

A.1/ This formulation assumes a constant marginal utility of income.

A.2/ The formulation (7) would provide a justification for the iso-elastic resource demand curve assumed often in interim reports 1 and 2.

A.3/ The assumption of multiplicative uncertainty is an essential assumption in the Diamond analysis. Because of that assumption, each firm can be viewed as producing a composite commodity; the theorem on the constrained Pareto Optimality of the market can be viewed as simply the statement that, given the set of "composite Commodities" being produced, if each firm acts as a price taker with respect to the price of the composite commodity it produces, then the market is Pareto Optimal. If stockholders (consumers) take the set of composite commodities produced as given and the price of each composit commodity as given as well, then clearly the only way that an action of the firm affects their welfare is through its affect on the valuation of the firm. Thus, there is unanimity among the shareholders both that the firm should maximize its stock market value, and that a particular level of investment will lead to the maximum stock market value. (For a more general discussion of the conditions for unanimity and the conditions under which the firm should maximize value is contained in Stiglitz [1976a] and Grossman and Stiglitz [1976].)

In the absence of multiplicative uncertainty, there may be disagreements both with respect to the effects of any given action on the market value and on the desirability of firm market value maixmization. Any given action of the firm will have an effect both on the value of the firm and on the consumption (or investment) opportunity set facing the individual even apart from the affect on the value of the firm, i.e. if there were two securities and three states of nature, the opportunity set is represented by a straight line in three dimensional space; with multiplicative uncertainty, individuals always perceive an action of the firm as moving the straight line out in a parallel manner (draw the plane defined by the line and the origin; then draw a straight line in that plane parallel to the original line) and all agree that such a movement is an improvement; but without multiplicative uncertainty the new opportunity set is a straight line which is not parallel to the original straight line. The change is preferred by some individuals, not by others. The relative importance of the consumption and valuation effects depends on the plans of the individuals with respect to the sale of securities. If, as in much of the recent literature, it is assumed as a condition of equilibrium than an individual neither plans to buy or sell shares, then the direct valuation effect is of no concern (since the price of anything which is neither bought nor sold has no affect, at the margin, on utility); but any reasonable model of the stock market involves trade, either because of life cycle effects, stochastic birth and death of firms, or differential information, and hence to some individuals but not all the valuation effect will be important; there will not be unanimity on the policy which the firm ought to pursue.

A.4/ The market valuation of the firm will in general depend on the debt-equity ratio of the firm, if there is a finite probability of bankruptcy; thus a change in the debt-equity ratio will have an affect both on the value of the firm and on the opportunity sets facing individuals. This is true even when there is marginal multiplicative uncertainty, i.e. increasing investment has a relative marginal effect on profits in different states of nature which is independent of the level of investment. In the absence of bankruptcy, with marginal multiplicative uncertainty again all individuals might agree on the objective of the firm, but with bankruptcy they will not. See Stiglitz [1970, 1975a].

A.5/ Hart [1975] provides an example of multiple equilibrium, in which in one of the equilibrium all individuals are better off than in another. This is an example of what I have called elsewhere a structural inefficiency (Stiglitz [1972]); such examples can occur even with a single commodity as I showed in Stiglitz [1972]. One of the Nash equilibria is Pareto Optimal, but there is no way of ensuring that this is the one which will occur.

On the other hand, the inefficiency which I notes in Stiglitz [1973] and which we discuss here is a marginal inefficiency; at the margin, the private market makes incorrect investment decisions. This would be the case even were there a unique equilibrium.

A.6/ In this analysis, we employ repeatedly the following approximations: We take all expansions around $\bar{Q}_2(\bar{Q}_d)$.

(25) $$p(Q) \approx p(\bar{Q}) + p'(Q - \bar{Q}) + \frac{p''}{2}(Q - \bar{Q})^2$$

(26) $$Ep(Q) \approx p(\bar{Q}) + \frac{p''\sigma_Q^2}{2}$$

(27) $$\sigma_p^2 \approx \frac{p''}{2}\sigma_Q^2$$

Hence

(28) $$p(Q) \approx \bar{p} + p'(Q - \bar{Q}) + \frac{p''}{2}[(Q - \bar{Q})^2 - \sigma_Q^2]$$

(29) $$V_Q = \{V_p + V_Y(\beta(\bar{Q}_1 - Q_1) + \gamma Q_s\}p' + V_Y p\gamma$$

(30) $$= -V_Y\{\Delta p' - p\gamma\}$$

(where we have used again the fact that $V_p = - Q_c V_Y$)

$$V_{QY} = -V_{YY}\{\Delta p' - p\gamma\} - V_Y mp'$$

(31) $$\frac{V_{QY}}{V_Y} = A[\Delta p' - p\gamma] - mp'$$

where

(32) $$A = \frac{-V_{YY}}{V_Y}$$

and

$$m = \frac{\partial Q_c}{\partial Y} \quad .$$

Thus letting

$$\xi \equiv \frac{Y^a}{Y^a + Y^b}$$

we can write

(33) $$\frac{V^a_{QY}}{V_Y} = -\frac{p}{Y}\left[\frac{1}{\xi}\left(\frac{\Delta}{Q} v - \gamma\right)R^a - v^a_c\right]$$

where

$$v \equiv -\frac{p'Q}{p}, \quad S^a_c = p\frac{Q^a_c}{Y^a}, \quad S = \frac{pQ}{Y} \quad .$$

Similarly

(34) $$V_Y \approx V_Y(\bar{Q}) + V_{YQ}(Q - \bar{Q}) + \frac{V_{YQQ}(Q - \bar{Q})^2}{2}$$

(35) $$EV_Y p = p(\bar{Q})V_Y(\bar{Q}) + \left[\frac{V_Y(\bar{Q})p''}{2} + \frac{p(\bar{Q})V_{YQQ}}{2} + p'V_{YQ}\right]\sigma^2_Q$$

(36) $$\frac{EV_Y p}{EV_Y} \approx p(\bar{Q}) + \left(\frac{p''}{2} + p' \frac{V_{YQ}}{V_Y}\right)\sigma^2_Q \quad .$$

If this is to be the same for all individuals V_{YQ}/V_Y must be the same.

$$pQ = p(\bar{Q})\bar{Q} + (p'\bar{Q} + p)(Q - \bar{Q}) + (\frac{p''Q}{2} + p')(Q - \bar{Q})^2 \tag{37}$$

$$\frac{EV_Y pQ}{EV_Y} \simeq p(\bar{Q})\bar{Q} + [\frac{p''Q}{2} + p' + (p'Q + p)\frac{V_{YQ}}{V_Y}]\sigma_Q^2 \quad . \tag{38}$$

Hence to a second order approximation, whenever

$$\frac{EV_Y p}{EV_Y}$$

is the same for all individuals, so is

$$\frac{EV_Y pQ}{EV_Y} \quad .$$

A.7/ Throughout this section, unless otherwise noted, all variables are those of the second period.

7. Imperfect Competition

7.1 The Behavior of a Monopolistic Resource Supplier

In principle the behavior of a profit-maximizing monopolist who controls the supply of an exhaustible resource is easily analyzed: if demand conditions permit, he will manipulate the supply so as to ensure that marginal revenue rises over time at a rate equal to the interest rate. The initial price (or quantity) will be chosen so that the total stock is just exhausted at some appropriate time. The implications of such behavior in the case of a constant-elasticity demand curve

$$Q = \gamma P^{\beta} \quad , \quad \gamma > 0 \quad , \quad \beta < 0$$

are well known. In this case,

$$\text{(1)} \qquad MR \equiv \frac{\partial(PQ)}{\partial Q} = (1 + \frac{1}{\beta})P \quad .$$

Cases where $\beta \geq -1$ are uninteresting, as they imply uniformly non-positive marginal revenue and the non-existence of a best policy for the monopolist. When $\beta < -1$,

$$\text{(2)} \qquad \frac{\dot{MR}}{MR} = r \Rightarrow \frac{\dot{P}}{P} = r$$

so that the conditions necessary for maximization of the present value of profits are also necessary for efficient intertemporal allocation. There may therefore be nothing wrong, in allocative terms, with a monopoly.

Obviously the constant elasticity function is a very special one, but it is a very convenient benchmark: in this role, we shall have reason to refer to it again later. In the meantime, we shall consider demand functions for which the elasticity varies in a simple and systematic way along the demand curve. It is clear that in reality we would expect the elasticity facing a monopolistic resource supplier to vary with price and quantity, though it is not entirely clear a priori what the nature of this variation should be.

One could for example argue that as the monopolist lowers his price and his market expands, this expansion brings him increasingly into competition with substitutes for his product. In the particular case of oil, the argument would run as follows: as the price is lowered, oil cuts increasingly into markets which by virtue of some particular technological or geographical characteristic had traditionally been the preserve of other fuels. (For example, electric utilities situated near coalfields might switch to oil firing.) In such cases, oil's advantage would be marginal and easily lost by small relative price changes. One would therefore expect the elasticity of demand facing oil producers to rise in absolute value as their price is lowered and their market share expands.

Although the above is plausible, there is an equally cogent argument leading to exactly the opposite conclusion. The above argument hinged on the existence of near-substitutes with which the resource-seller comes increasingly into competition as the price is lowered: the alternative case rests on the fact that as the price of the resource is raised, this increases the incentive to invent substitutes that did not previously exist, or to proceed with development work on potential substitutes whose development has been held in abeyance while the resource price has been low. Again illustrating the general with the particular case of oil, the argument would run in terms of higher oil prices encouraging the development of shale oil, of dual-firing systems in boilers, and a variety of other developments which would increase the elasticity facing oil sellers. According to this argument, then, the elasticity facing a resource-supplier is likely to be an increasing, rather than a decreasing, function of his price.

We have established that it would be sensible to consider the behavior of a monopolist facing a demand curve with a variable elasticity, though as mentioned earlier it does not seem possible to determine a priori what the nature of this variation should be. It will however emerge in due course that a monopolist will behave differently in the various possible different cases, so that it is important in policy terms to differentiate between them. We therefore present below a number of demand functions for which the elasticity is a simple function of output, increasing, decreasing, (or showing both properties). The demand functions involve only two parameters, and should in principle make it easy to conduct empirical work designed to discriminate between the cases already mentioned. The first pair of functions we shall consider is

$$(3) \qquad P = \gamma e^{-\beta Q} \qquad \gamma , \beta > 0$$

and

$$(4) \qquad P = \gamma e^{\frac{1}{\beta+Q}} \qquad \gamma , \beta > 0 \quad .$$

If η is the elasticity of demand,[1/] it is easily shown that for function (3)

(5) $$\eta(Q) = -\frac{1}{\beta Q}$$

and for function (4)

(6) $$\eta(Q) = -\{\frac{\beta^2}{Q} + 2\beta + Q\} .$$

In addition, for form (3) marginal revenue is given by

(7) $$M.R. = P[1 - \beta Q]$$

and for function (4)

(8) $$M.R. = P[1 - \frac{Q}{(\beta + Q)^2}] .$$

It is convenient to catalogue the properties of these functions.

7.1.1 Case I. $P = \gamma e^{-\beta Q}$: $\eta(Q) = -1/\beta Q$, increasing in absolute value as P rises and Q falls.

Marginal revenue is zero for $Q = 1/\beta$, positive for $Q < 1/\beta$ and negative for $Q > 1/\beta$ (see figure

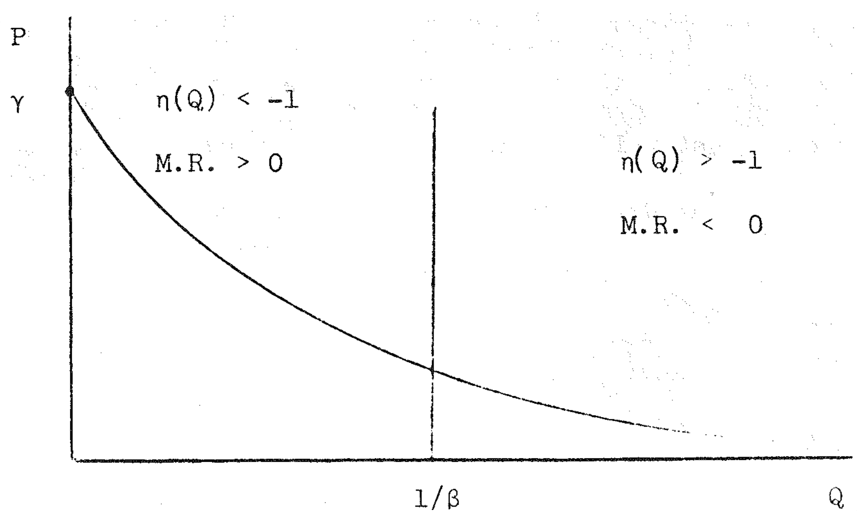


Figure 1.

There is a "choke price" of γ at which demand falls to zero. This can be interpreted as a price sufficiently high to cause users to introduce substitutes. As quantity rises, price falls to zero.

7.1.2 Case II. $P = \gamma e^{\frac{1}{\beta+Q}}$: We shall restrict our attention to cases where $\beta \geq 1/2$. We have $\eta = -\{(\beta^2/Q) + 2\beta + Q\}$, so that η is minus infinity when $Q = 0$ or $Q = +\infty$, and reaches a maximum of -4β when $Q = \beta$. There is a "choke price" $\gamma e^{1/\beta}$ at which demand falls to zero, and a limit price of $\gamma(< \gamma e^{1/\beta})$ to which price falls as output expands indefinitely. Marginal revenue, given by

$$M.R. = \gamma e^{\frac{1}{\beta+Q}}[1 - \frac{Q}{(\beta + Q)^2}],$$

falls monotonically from $\gamma e^{1/\beta}$ to γ as Q rises. This demand function thus exhibits both of the characteristics referred to earlier: there is an upper limit to price, set presumably by the cost of bringing in substitutes, and the elasticity rises as this is approached. In addition, reduction of price and expansion of the market also lead to increases in the elasticity, presumably as more intense competition is encountered.

It can be seen that these two demand functions model very naturally the cases of variable elasticity referred to earlier: in either case the parameter γ has a natural interpretation in terms of the price of substitutes or a "trigger price" at which substitutes are introduced, and β is a parameter which controls the rate of change of the elasticity along the curve.

We shall analyze the optimal behavior of a monopolist in the two cases, and use the results as a basis for evaluating various policy resources designed to change demand conditions facing suppliers. We shall also present an alternative parameterization of demand conditions which will allow the analysis to be conducted in slightly different terms.

<u>The Monopolist in Case I, $P = \gamma e^{-\beta Q}$:</u> Marginal revenue decreases with output, being negative for $Q \geqq 1/\beta$. A monopolist will thus ensure that $Q_t < 1/\beta$ for all t, and will supply a quantity that falls over time, causing price to rise. But from the facts that

$$M.R. = P(1 + \frac{1}{\eta})$$

and

$$\frac{\dot{MR}}{MR} = r = \frac{\dot{P}}{P} + \frac{\dot{\overline{(1 + \frac{1}{\eta})}}}{(1 + \frac{1}{\eta})}$$

it is easily verified that $\dot{P}/P < r$. Price rises less rapidly than along an efficient path, so that there is excessive conservation.

<u>The Monopolist in Case II, $P = \gamma e^{\frac{1}{\beta+Q}}$:</u> Figure 2 shows the relevant behavioral relations:

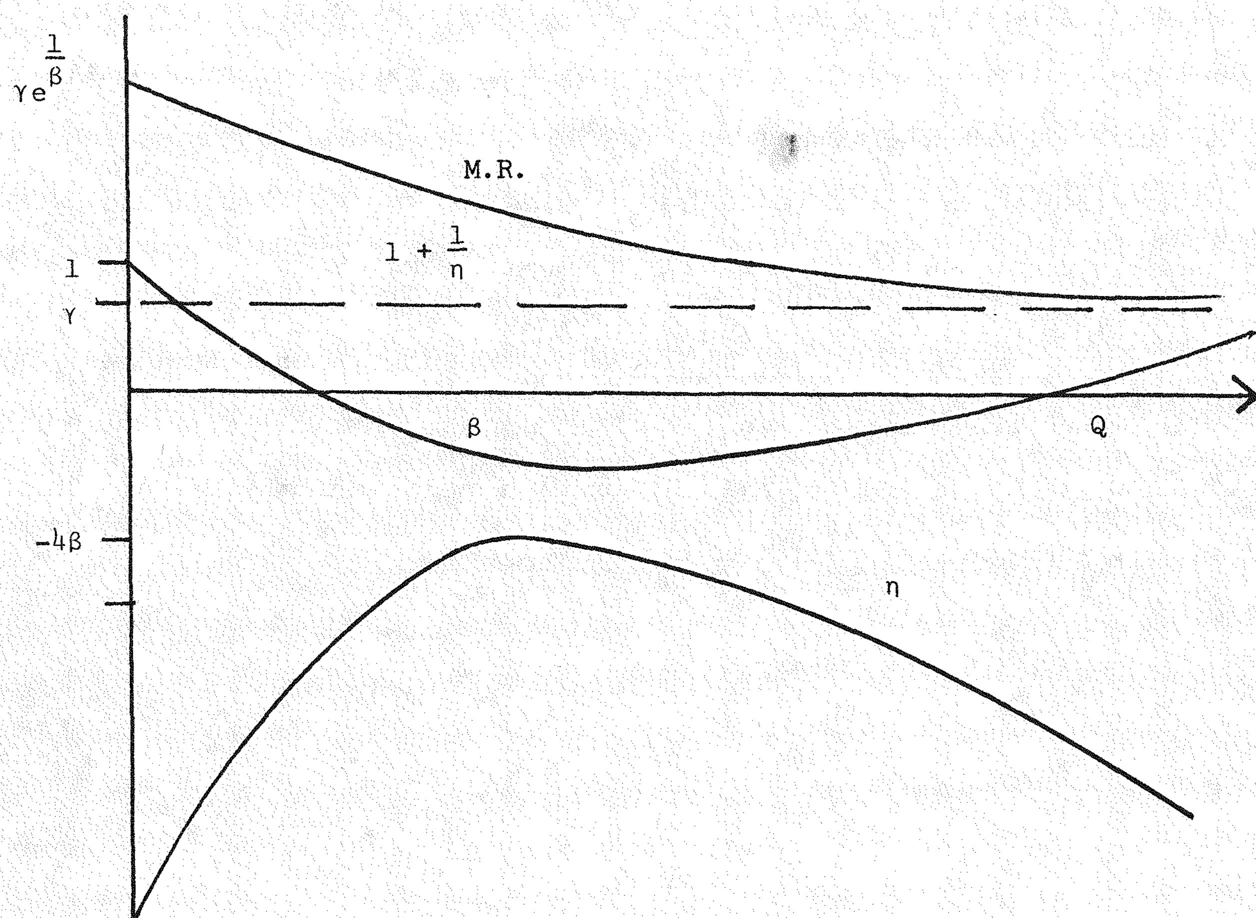


Figure 2.

The monopolist will seek to make marginal revenue rise over time at the rate of interest r, and this requires falling quantity and rising price. It is readily verified that for $Q > \beta$, so that $(1 + 1/\eta)$ is falling, the monopolist would wish to have $\dot{P}/P > r$: at $Q = \beta$, $\dot{P}/P = r$, and for $Q < \beta$, $\dot{P}/P < r$. Now it seems likely that a monopolist will not in fact be able to sustain a policy along which $\dot{P}/P > r$: stocks of the resource will offer super-normal rates of return, consumers will wish to buy unlimited quantities, and the market will not clear. If one accepts this argument, then the monopolist is constrained to have $\dot{P}/P \leq r$, in which case $\dot{P}/P = r$ for $Q \geq \beta$ and $\dot{P}/P < r$ for $Q < \beta$. There is once again excessive conservation.

General Comments on the Case of $|\eta|$ Increasing with Q

It is worth observing that, whenever a demand function has the property that the absolute value of the elasticity of demand rises with output, the resulting behavior of a monopolist is in some sense extreme. In such a case there are two possibilities: (1) If marginal revenue is monotonically declining with output the monopolist will wish to set $\dot{P}/P > r$, and will be constrained by the effects discussed above. (2) The alternative case is where marginal revenue rises with output (as for the function $P = \gamma e^{1/\beta Q}$ for which $\eta = -\beta Q$ and M.R. $= \gamma e^{1/\beta Q}\{1 - 1/\beta Q\}$), the equation of M.R. to an exponentially-rising shadow price produces a minimum rather than a maximum of the relevant Hamiltonian, and optimal behavior involves setting output to its upper or lower bound.

7.1.3 The Effects of Parameter Changes

The demand functions presented are characterized by two parameters, γ and β. We shall argue below that the effects of many policy measures designed to promote conservation or to increase the flexibility of energy-using equipment can be interpreted as changes in γ or in β, or as simple combinations of the two. In this section we therefore analyze the effects of such parameter changes in some detail.

An increase in γ shifts the demand curves to the right, raising the price associated with any given sales volume. The elasticity at any given

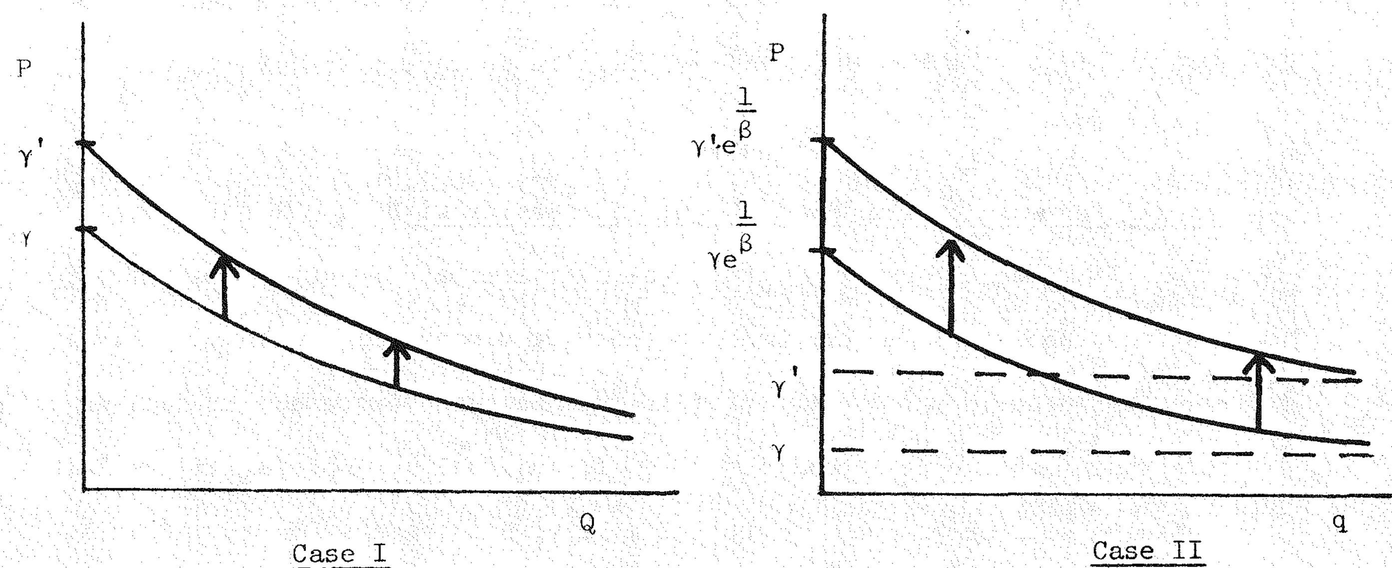


Figure 3

quantity is unaltered (at $-1/\beta Q$ or $-(\beta + Q)^2/Q$), though of course the elasticity at any given price does change. In terms of prices, the elasticities are

$$\eta(P) = -\frac{1}{\log\left(\frac{\gamma}{P}\right)} \qquad \text{Case I}$$

$$\eta(P) = -\frac{\gamma}{\left(\log\frac{P}{\gamma}\right)^2\left(\log\frac{P}{\gamma} - \beta\right)} \qquad \text{Case II}$$

so that in Case I (elasticity falling with market share) an increase in γ lowers the elasticity in absolute value at any given price level, and in Case II the reverse happens. It is also obvious that in Case I an increase in γ raises the upper bound on price (the "price of a substitute") and in Case II raises this and the lower bound (the "competitive price"), as shown in figure 3.

In addition, we can see that in Case I the requirement that M.R. ≥ 0 imposes an effective lower bound of γ/e (price when $Q = 1/\beta$) on price, whereas there is no such limitation in Case II -- though the price is of course bounded above and below by functions of γ. Table 1 summarizes the effects of an incrase in γ.

Table 1

Effects of an Increase in γ

	Case I	Case II
Lower bound on price	Rises ($= \gamma/e$)	Rises ($= \gamma$)
Upper bound on price	Rises ($= \gamma$)	Rises ($= \gamma e^{\frac{1}{\beta}}$)
Elasticity at given Q	Unaltered	Unaltered
Elasticity at given P	Decreases	Increases
Price at given Q	Increases	Increases
Q at which M.R. changes sign	Unaltered	Not applicable

The effects of a change in β are more complex. In Case I,

$$P = e^{-\beta Q} \quad , \quad \eta(Q) = -\frac{1}{\beta Q}$$

an increase in β lowers the absolute value of the elasticity at any given sales volume, and also lowers the rate at which the elasticity rises as the price rises. The elasticity at any given price in unchanged, and the demand curve is shifted to the left, this time with its left hand end point $(Q = 0, P = \gamma)$ unchanged (figure 4).

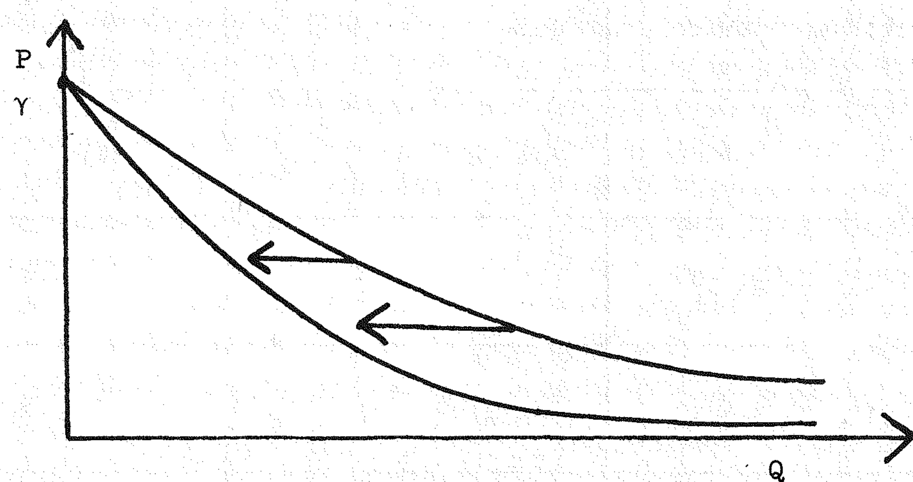


Figure 4

The marginal revenue curve is also shifted to the left:

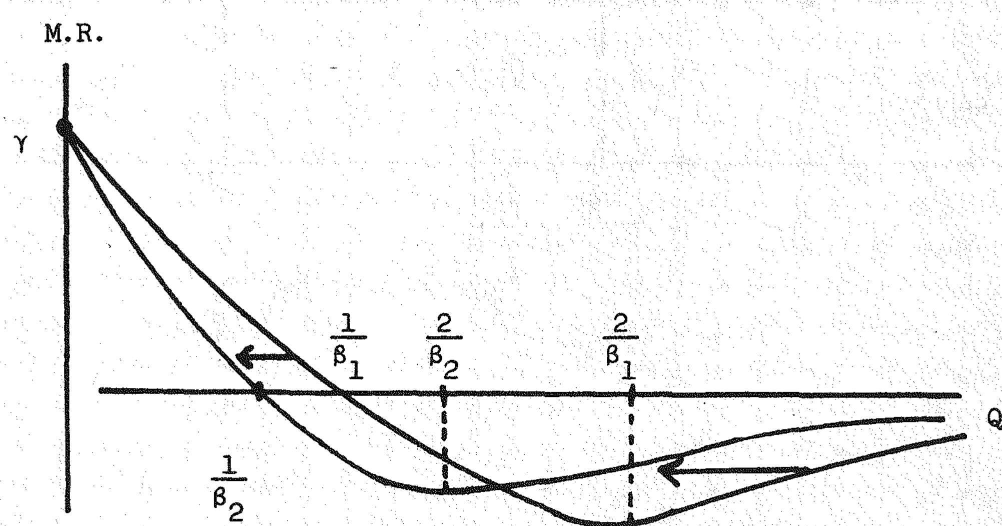


Figure 5

In Case II,

$$P = \gamma e^{\frac{1}{\beta+Q}} \quad , \quad \eta(Q) = -\{\frac{\beta^2}{Q} + 2\beta + Q\}$$

an increase in β raises the elasticity at any given sales volume. It also shifts the curve relating η to Q as shown in figure 6:

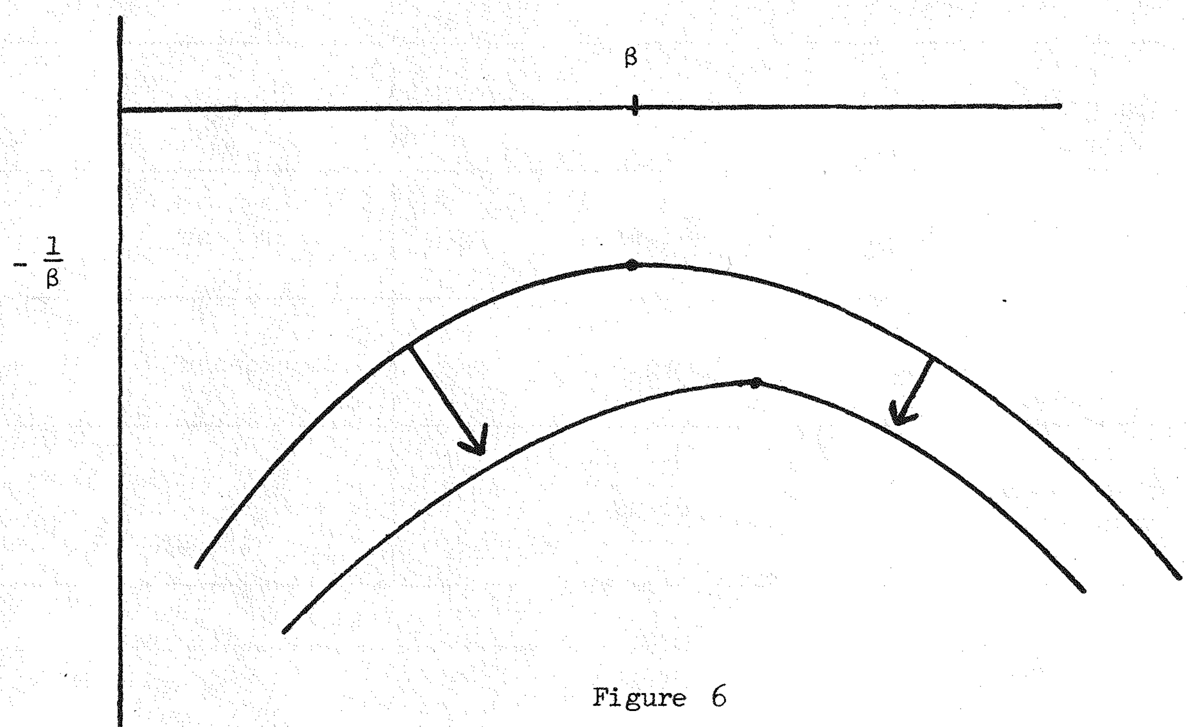


Figure 6

The elasticity at any given price increases, the "price of a substitute" $\gamma e^{1/\beta}$ falls, and the demand curve is shifted to the left.

Again, these results can conveniently be summarized in Table 2.

Table 2

Effects of an Increase in β

	Case I	Case II
Lower bound on price	Unaltered $(= \frac{\gamma}{e})$	Unaltered $(= \gamma)$
Upper bound on price	Unaltered $(= \gamma)$	Decreases $(= \gamma e^{\frac{1}{\beta}})$
Elasticity at given Q	Decreases	Increases
Elasticity at given P	Unaltered	Increases
Price at given Q	Decreases	Decreases
Q at which M.R. changes sign	Decreases	Not applicable

Before moving on to an evaluation of policy measures, we shall for completeness consider the case of the constant-elasticity demand function

$$Q = \gamma P^{\beta} \quad , \quad \gamma > 0 \quad , \quad \beta < 0 \quad .$$

Clearly an increase in γ shifts this to the right: a decrease in β raises the absolute value of the constant elasticity but leaves the quantity sold unchanged at γ when $P = 1$. Decreasing β thus pivots the demand curve about $(\gamma,1)$:

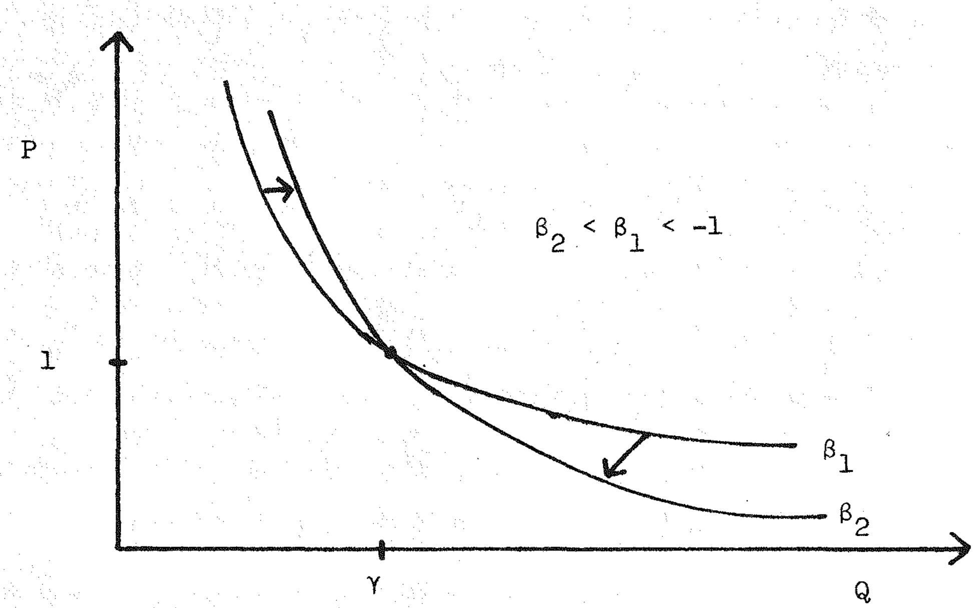


Figure 7

7.1.4 Evaluation of Policy Measures

We shall conduct the evaluation of policy measures in two stages:

(i) Evaluate the desireability or otherwise of changes in γ and β, and

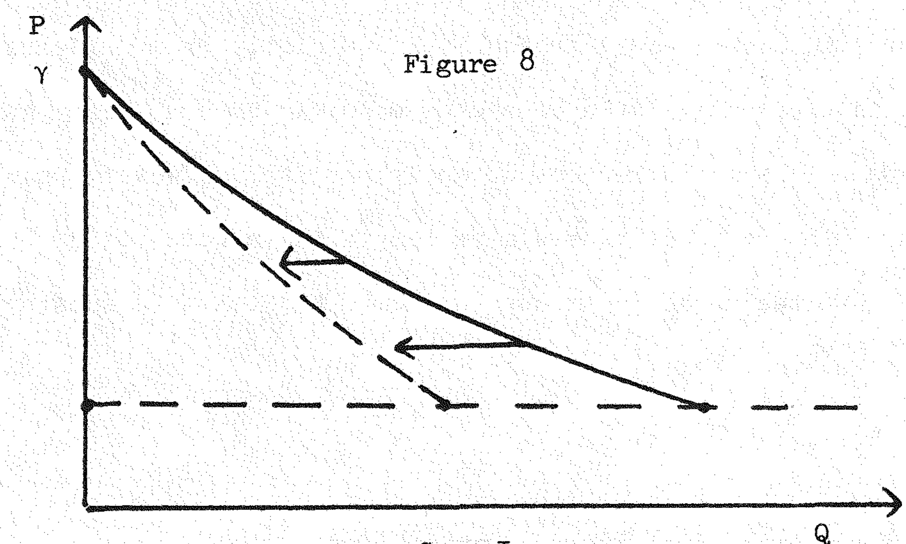
(ii) Identify various policy proposals with changes in γ and β, allowing them to be brought within the compass of conclusions in (i).

Examination of Tables 1 and 2 reveals that the following statements can be made:

An increase (decrease) in γ raises (lowers) the interval within which price muse move, and shifts the demand curve to the right (left). It leaves unchanged the interval within which quantities will move.

An increase (decrease) in β leaves unchanged the interval within which prices will move and shifts the demand curve to the left in Case I. In Case II, the demand curve is moved similarly but the interval within which price moves is shifted downwards. An increase (decrease) in β lowers (raises) the interval within which quantities will move in Case I, but not in Case II.

An obvious consequence of this is that we can describe an increase (decrease) in γ as harmful (beneficial) as it raises (lowers) the price to be paid at any quantity, without providing any incentive to the supplier to raise (lower) the interval within which quantity moves. In Case I, the evaluation of a change in β is less clear. An increase lowers the price paid for any given quantity, which is ceteris paribus beneficial. But it also lowers the interval within which quantity must move (see figure 4 which shows the changes in the operative (M.R. ≥ 0) regions of the demand curves). This is potentially harmful, though one can show that if the effect of the increase in β is to either lower or leave unchanged the price at which the monopolist initially supplies the resource, the average price paid by consumers over the duration of the monopoly will be lower at higher β. Although we have been unable to establish this definitively, this seems to be the likely case. We therefore conclude that an increase in β is beneficial, and a decrease harmful, in Case I.



Case I

Increase in β

In Case II, matters are more straightforward. An increase in β lowers the price at any output level and lowers the upper bound on price, leaving the lower bound unaltered. This must be beneficial from the consumer's viewpoint.

Having evaluated the desirability or otherwise of changes in γ and β, we can now pass to the next stage of the analysis and attempt to link specific policy proposals to certain patterns of changes in these parameters. It should be noted that the analysis is intended to be illustrative of the methodology, rather than comprehensive. It should also be noted that we are concerned only to evaluate the sign of gross benefits to the U.S.: no attempt is made to cost the alternative programs or to consider their world-wide welfare effects.

A Tax on Oil

Clearly this lowers the relative price of substitutes, lowering the upper or lower bounds on the price of oil and reducing the quantity demanded at any price. It is immediately apparent from Table 1 that this is equivalent to a decrease in γ: it will therefore decrease the elasticity at any price in Case I and increase it in Case II. From our earlier arguments, it is clear that for a bilateral monopoly such a policy has a beneficial impact in that it reduces monopoly profits. However, unless the tax is carefully designed, the distortion in the allocation of oil may be increased. Furthermore, in an open economy, the tax is a subsidy to consumer groups that are not taxed.

Encouragement of Conservation Measures

In this case the costs of the substitutes for any particular fuel are presumably unchanged, but demand is lower at any price. Table 2 makes it clear that in Case I this can be modelled as an increase in β, and so it is on balance beneficial in its impact on import prices. In Case II, such a move does not correspond exactly to changes in γ or β, though is best approximated by an increase in β. To model such a change exactly in Case II it would be necessary to add a third parameter σ to the functional form, making it $P = \gamma e^{\sigma/\beta+Q}$.

Convertibility Subsidies

Under this heading we have in mind subsidies embodied in preferential tax treatment on the installation of equipment which may burn either of several energy sources. The primary effect of such a move would be to increase the elasticity of demand for a particular fuel (say oil) at any given price, and one might reasonably assume that even if quantity were held constant, elasticity would increase. It is not entirely clear what the effect of such a move would be on the limits to price, but if the flexibility of equipment is increased there will presumably be a reduction in the upper bound on price. Inspection of Tables 1 and 2 shows that in Case I such a combination of effects is best modelled by decreases in both γ and β. From our earlier analysis, the former is unambiguously beneficial, and the latter harmful. This makes the net benefit hard to evaluate: however, as the main effect is probably to raise the elasticity at any given price (lower γ), and as a decrease in γ is more clearly beneficial than a decrease in β is harmful, it seems that on balance convertibility subsidies will have a beneficial effect in Case I. Such a conclusion should be tempered by the realization that a subsidy might, by lowering the average fixed cost of energy use, raise the demand for energy in general. Such an effect might need to be offset by a general tax on energy use.

In Case II, matters are still not clear cut. The effect of the subsidy is now best modelled as an increase in γ and β, with the former now harmful and the latter beneficial. The only clue to the net effect comes from the following argument. We expect that the upper limit on price, $\gamma e^{1/\beta}$, will fall. This requirement imposes a lower bound on the increase in β relative to γ -- to be exact, if γ and β change respectively from γ_1 to γ_2 and β_1 to β_2,

$$\beta_2 > \frac{1}{\left(\frac{1}{\beta_1} - \log \frac{\gamma_2}{\gamma_1}\right)} .$$

The fact that the change in β is bounded below by an increasing function of the change in γ at least suggests, though no more, that there will be cases where the β-effect predominates and the net effect is beneficial -- subject, of course, to the qualification already made about the effect of the subsidy on demand.

It is also interesting to consider the impact of convertibility subsidies in the context of a constant-elasticity demand function

$$Q = \gamma P^{\beta} \quad , \quad \gamma > 0 \quad , \quad \beta < 0 \quad .$$

They are clearly best modelled as a decrease in β, which pivots the demand curve clockwise about the point $Q = \gamma$, $P = 1$. For quantities less than γ, price is raised, and vice versa: the overall impact therefore depends on the size of the monopolist's stock, being harmful when this is less than some critical level, but possibly beneficial if it is large enough.

7.1.5 Concluding Comments on Policy Options and Demand Parameterizations

As stressed earlier, the above is not intended as a comprehensive evaluation of policy options, but rather as an illustration of the methodology that could be based on the parameterization of demand functions presented in the earlier sections. Obviously other parameterizations could be developed, and indeed we have on occasions seen the need for a function of the form

$$P = \gamma e^{\frac{\sigma}{\beta+Q}} \quad .$$

Here the extra parameter σ would give a further degree of freedom in shifting the function to model certain effects. However, the gain in generality is spurious, as one can easily show that all three parameters of this equation could not be separately identified in empirical work.

An alternative parameterization which merits consideration is the following:

$$P = (Q_0 + Q)^{-\beta} \qquad \text{Case III}$$

$$P = \left(\frac{1}{Q_0} + \frac{1}{Q}\right)^{\beta} \qquad \text{Case IV}$$

For Case III:

$$\eta = -\frac{1}{\beta}\left\{\frac{Q_0}{Q} + 1\right\}$$

$$M.R. = \frac{1}{(Q_0 + Q)^{\beta}}\left\{1 - \frac{\beta}{\frac{Q_0}{Q} + 1}\right\}$$

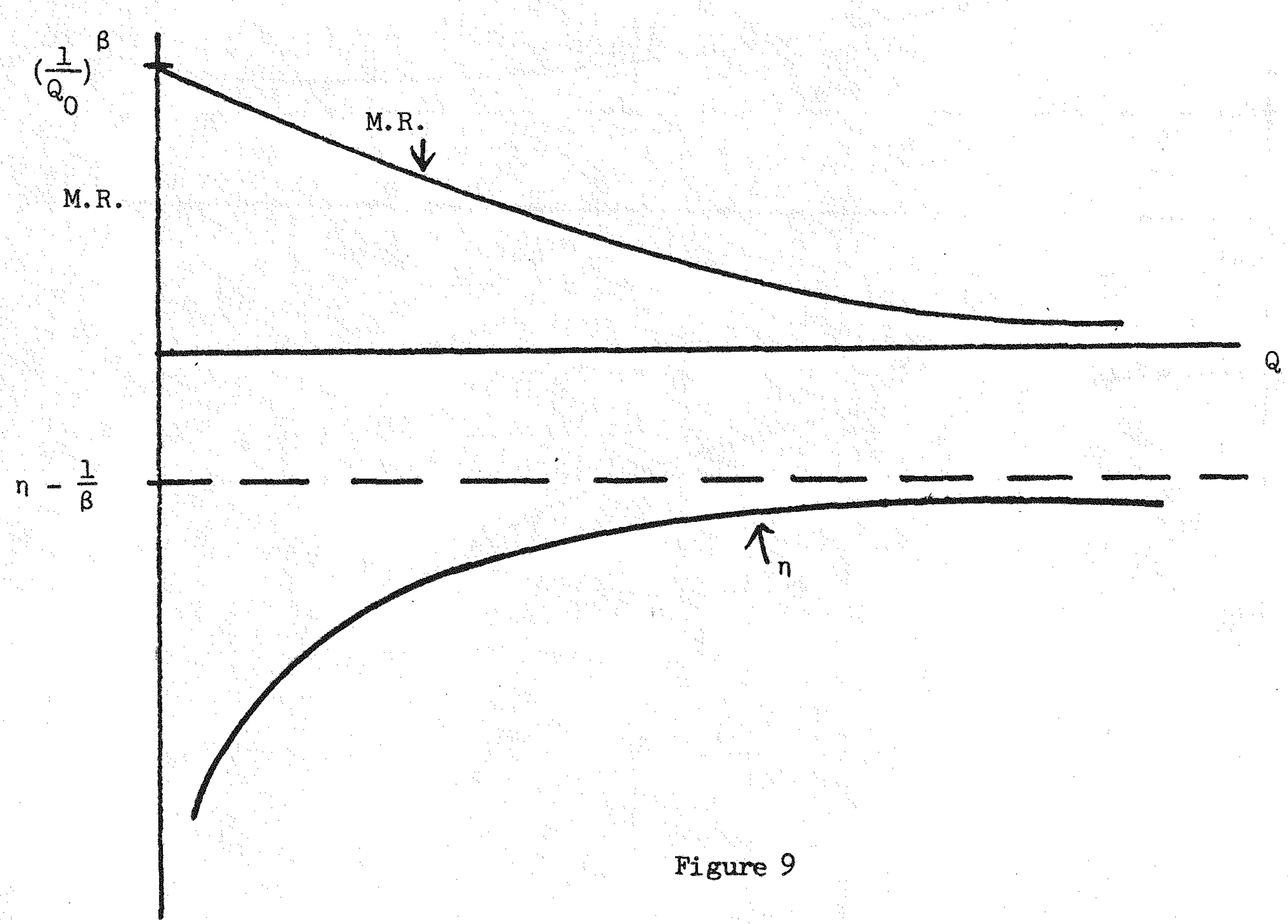


Figure 9

We have here a 2-parameter function whose behavior parallels that of Case I closely, except that marginal revenue is positive at all output levels.

For Case IV:

$$\eta = -\frac{1}{\beta}\{\frac{Q}{Q_0} + 1\}$$

$$M.R. = (\frac{1}{Q_0} + \frac{1}{Q})^{\beta}\{1 - \frac{\beta}{\frac{Q}{Q_0} + 1}\}$$

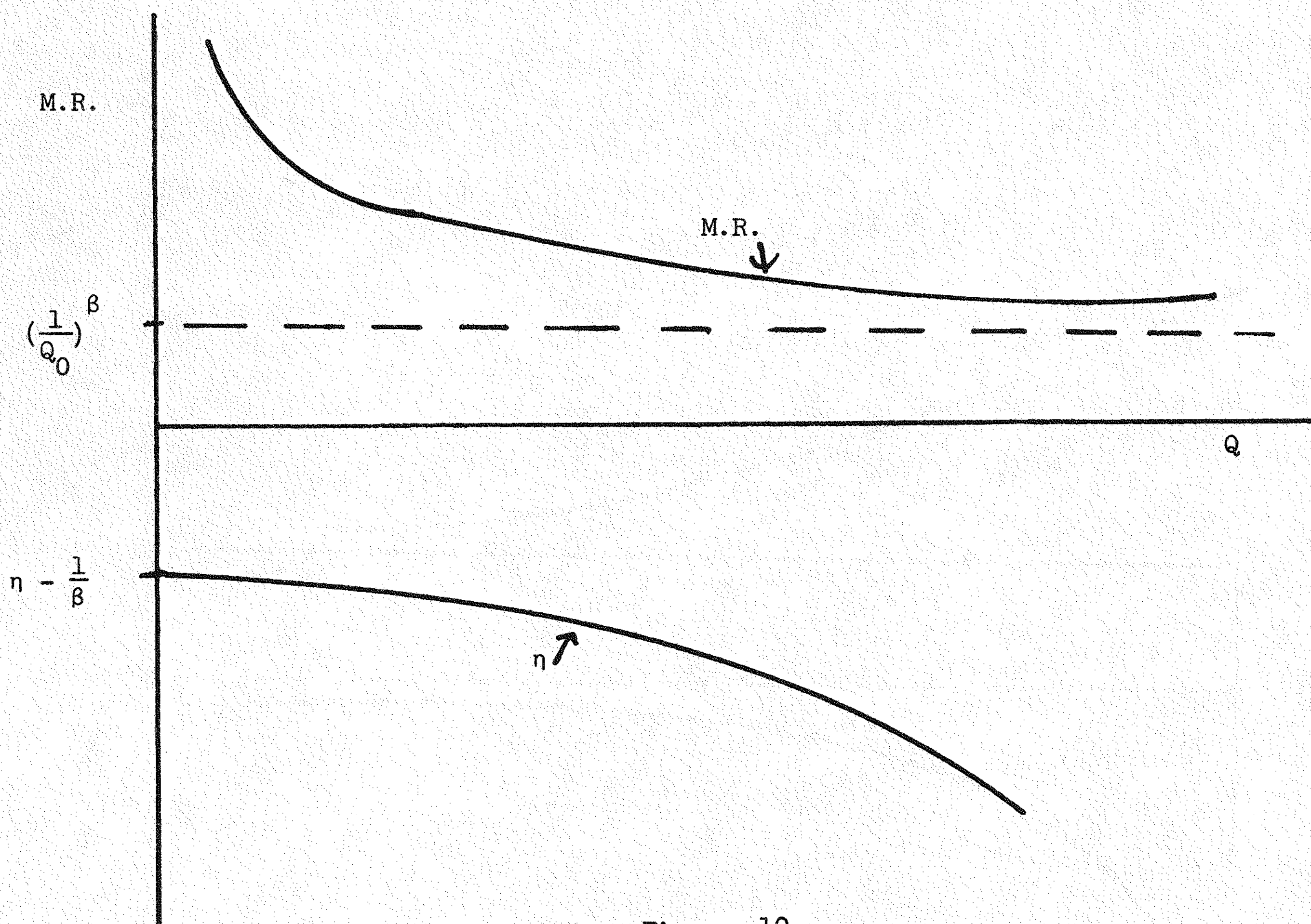


Figure 10

This function models a case not separately identified by Cases I and II, viz. the case where the elasticity increases uniformly in absolute value along the demand curve. In keeping with our earlier comments on such cases, one can verify that along the monopolists optimal depletion policy, the constraint $\dot{P}/P \leq r$ will be binding.

It is clear that much of our earlier analysis could have been presented in terms of the functions of cases III and IV, and one can readily verify that few conclusions would differ: the choice is largely a matter of convenience, which will depend on the particular problem in hand.

7.2 The Profit-Maximizing Cartel

The preceding section discussed the profit-maximizing production allocation determined by a monopolist on the supply of a resource. This is, of course, an extreme case. While some resource markets are dominated by a particular producer or collection of producers (for example, the OPEC cartel), there are existing and potential suppliers that act as price-takers. This collection of competitive or socially-managed producers we term the competitive fringe. The set of producers who explicitly consider the effects of their coordinated supply decisions on the resource price we term the cartel.

In this section we show that the supply response of the competitive fringe can significantly limit the monopoly power of the cartel. The effect of the fringe on the cartel allocation depends on such factors as the size of the stock owned by the fringe, constraints on the rate of extraction, and the cost of extraction. The results suggest that in a market dominated by a cartel, the existence of a fringe leads to allocations that differ dramatically from those that characterize a monopoly on the supply of a resource.

The following notation is used in this chapter.

$S_{\ell}(t)$: Total exhaustible resource stock owned by the cartel at time t.

$S_f(t)$: Total exhaustible resource stock owned by the competitive fringe.

$k_f(t)$: Output capacity of the fringe.

$Q(t)$: Total exhaustible resource output (assumed equal to consumption unless stated otherwise).

$P(Q,t)$: Demand price of the resource.

$\alpha = -\frac{P(Q,t)}{Q(t)} \frac{\partial Q(P,t)}{\partial P}$: Magnitude of demand elasticity.

$q_\ell(t)$: Cartel output.

$R_\ell(t)$: Cartel revenue.

$MR_\ell(t)$: Cartel marginal revenue.

We assume in this section that a perfect substitute is available at price $\bar{P}$ with zero development cost. For simplicity, we assume the extraction cost of the resource is zero (up to the capacity constraint), and the market rate of interest is a constant, r.

We have argued that the competitive sector will hold back production if expected capital gains on the exhaustible resource exceed the market rate of interest, and produce at capacity if expected capital gains fall short of the rate of interest. The competitive sector will supply at the rate demanded if capital gains equal the interest rate, or it

$$\frac{\dot{P}(t)}{P(t)} = r \quad .$$

An analogous statement applies to the cartel, but with price replaced by marginal revenue. Cartel production will equal demand if

$$\frac{\dot{\overline{MR}}(t)}{MR(t)} = r \quad ,$$

where $MR(t) = P(t)[1 - 1/\alpha]$.

7.2.1 Constant Demand Elasticity $(\alpha > 1 \; ; \; \dot{\alpha} = 0)$

For this case, since extraction cost is assumed to be zero,

$$\frac{\dot{\overline{MR}}(t)}{MR(t)} = \frac{\dot{P}(t)}{P(t)} = r \quad .$$

Both the cartel and the competitive sector face the same rate of return at each point in time, and the same terminal condition: namely,

$$P(Q(T)) = \bar{P} \text{ when } S_{\ell}(T) = S_f(T) = 0 \quad .$$

The price rises at the rate of interest until $P(t) = \bar{P}$, and then remains at $\bar{P}$ (figure 11).

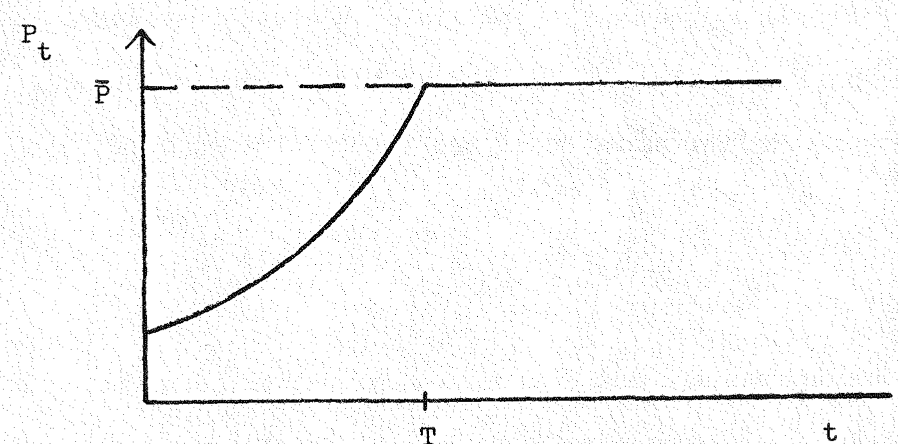


Figure 11

The effect on the allocation of the resource would be the same if either the fringe or the cartel's output were constrained by production capacity. The constrained sector would produce at capacity until demand fell below the capacity constraint. If

$$P(q_i(k)) > \bar{P}\ ,\ i = \ell, f$$

the resulting allocation would appear as shown in figure 11 except that both the substitute and the resource would be produced for $t > T$. If otherwise, the time trajectory of prices might appear as shown in figure 12. Of course, if capacity is a binding constraint, there is an incentive to invest in additional capacity.

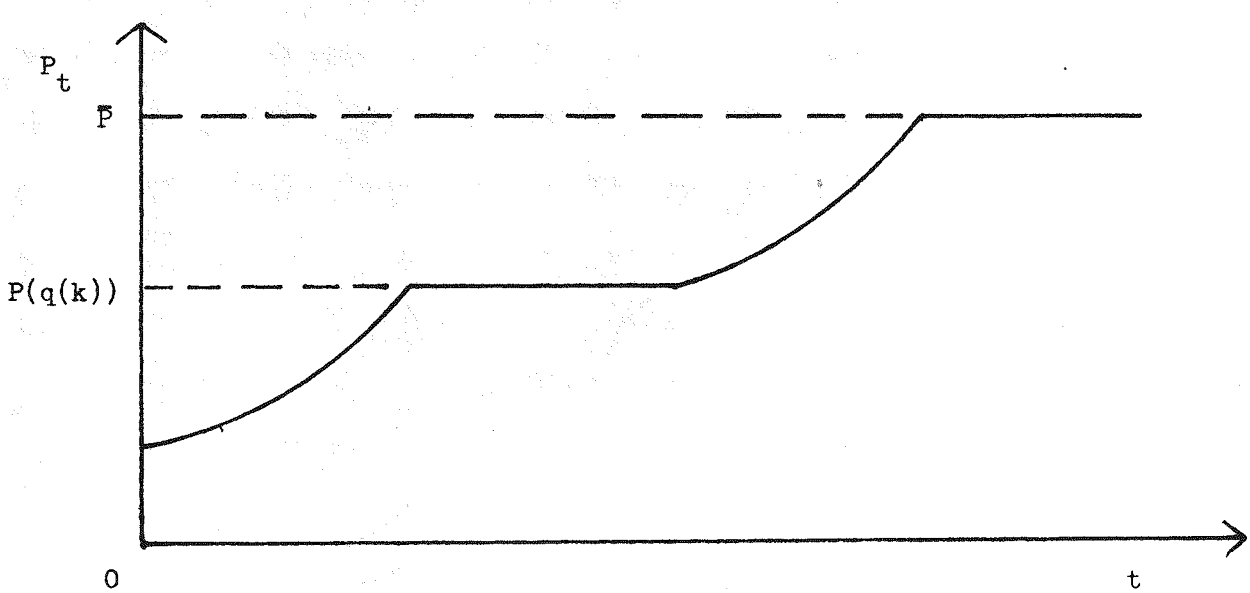


Figure 12

The important observation, however, is that for this case, where $\alpha > 1$ and $\dot{\alpha} = 0$, there is no asymmetry in the supply responses of the competitive and monopolistic sector. The cartel is in name only, and has no distortionary effect on the aggregate allocation of the exhaustible resource.[2/]

While an interesting example, this is not necessarily the central case. There are finite extraction costs to consider. Perhaps more important, the demand elasticity may vary over time, and empirical studies have suggested that in recent years the magnitude of the short-run demand elasticity for oil has been signicantly less than one.[3/]

.2.2 Elasticity greater than unity (in magnitude) and increasing over time ($\alpha > 1$; $\dot{\alpha} > 0$)

In the first interim report, we gave a specific example comparing the socially optimal and pure monopoly allocations when demand elasticity was above unity and increasing, and it was shown that the monopolist was over-conservative. The result is repeated in figure 13.[4/]

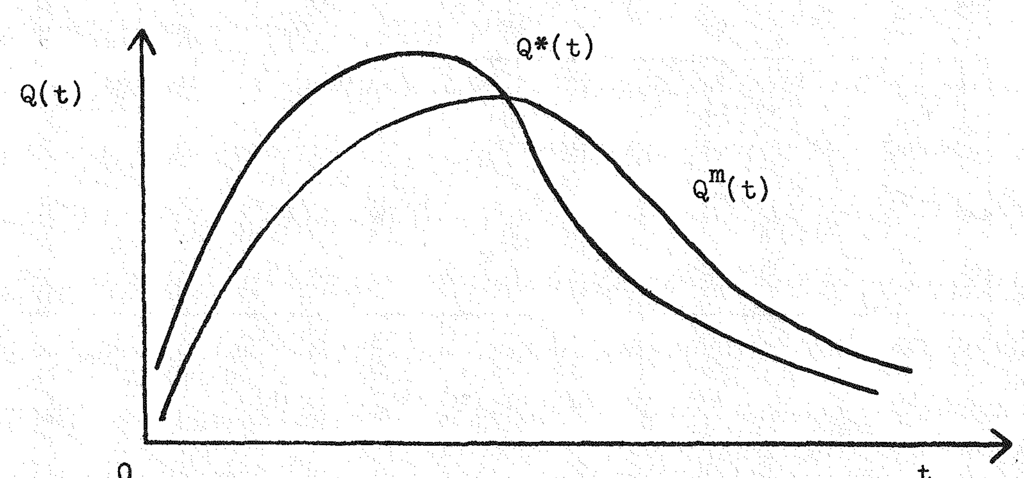


Figure 13

What we shall explore now is the influence of the competitive fringe on the total resource allocation.

Equilibrium rates of return for the competitive sector and the cartel are

$$\frac{\dot{P}(Q)}{P(Q)} = r \qquad \text{(competitive sector)}$$

$$\frac{\dot{P}(Q)}{P(Q)} = r - \frac{\dot{\phi}}{\phi} \qquad \text{(cartel)}$$ [5/]

where

$$\phi(t) = 1 - \frac{1}{\alpha(t)} ,$$

and

$$\frac{\dot{\phi}}{\phi} > 0 \text{ if } \dot{\alpha} > 0 .$$

Assume no constraints on production capacity. In this case, the supply response of the competitive sector will constrain the price so that

$$\frac{\dot{P}}{P} = r$$

if $S_f(t) > 0$. If $\dot{\phi} > 0$, the cartel's rate of return on its resource stock exceeds the market rate of return, and the cartel should postpone extraction to maximize profits.

If $\ddot{\phi} < 0$ (which would be the case if $\ddot{\alpha}(t) < 0$), the distortion in the cartel's allocation would be lessened (relative to the socially optimal norm) by the postponement. The fringe in this case takes away some monopoly rent by forcing the cartel to postpone extraction. It can do this because the absence of any capacity constraint will insure a competitive rate of return. The result would be as shown in figure 14.

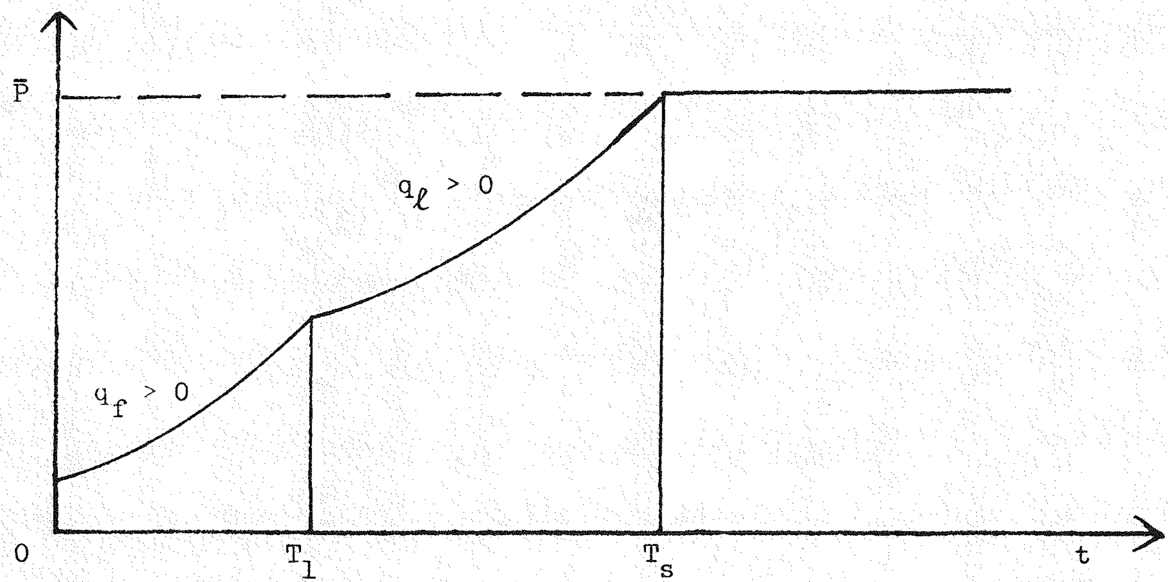


Figure 14

For

$$t \in (0,T_1) \quad , \quad \frac{\dot{P}}{P} = r$$

and for

$$t \in (T_1,T_s) \quad , \quad \frac{\dot{P}}{P} = r - \frac{\dot{\phi}}{\phi} \ .$$

The effect of the fringe on the monopoly power of the cartel depends on $\alpha(t)$ and S_f. Note that the allocation would be efficient if $\dot{\phi} = 0$ for $t \geq T_1$.

If the production capacity of the fringe is limited, the cartel can exercise monopoly discretion from the initial time period. Furthermore, the result is not symmetric if it is the cartel that is capacity limited and not the fringe.

Cases 3.2.1 and 3.2.2 are offered for illustration, and certainly other examples can be developed, but with diminishing returns. We now consider the case of imperfect competition when $\alpha < 1$, which in view of empirical results would appear to be the central category.

7.2.3 Optimal cartel pricing when demand is inelastic ($\alpha < 1$)

$\alpha < 1$; no capacity constraint

When $\alpha \leq 1$, a monopolist's optimal (constant price) strategy is to set the market price of the resource just below the price of the substitute source of supply. 6/ The demand correspondence for the monopolist may look as shown in figure 15

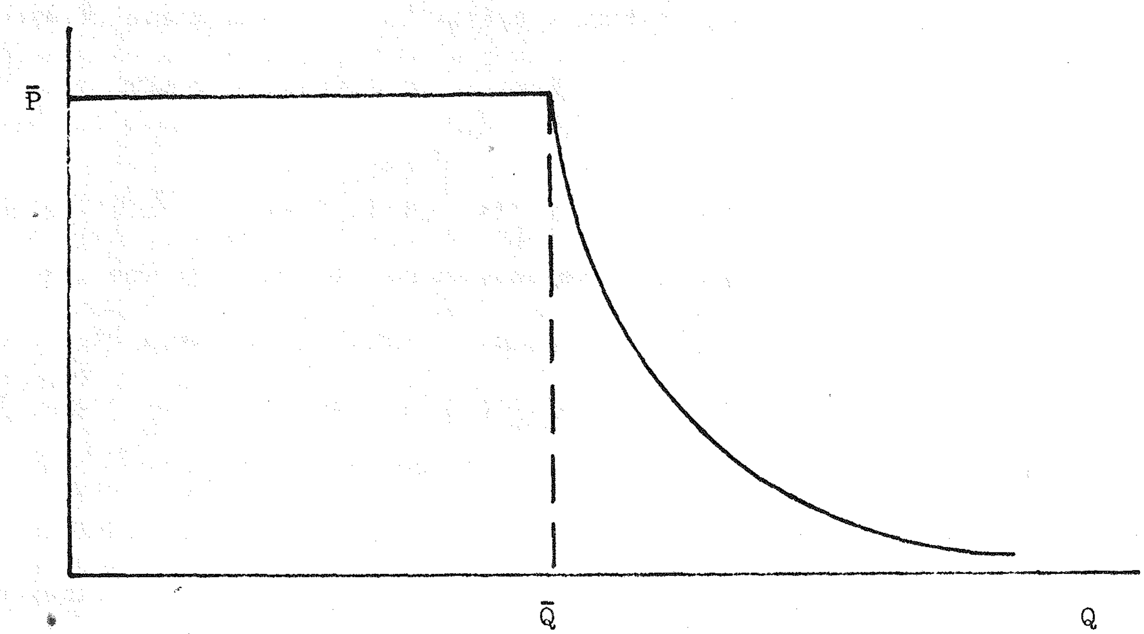


Figure 15

For $Q > \bar{Q}$, by assumption.

$$MR_{\ell} = P(Q)[1 - \frac{1}{\alpha(Q)}] < 0 \quad .$$

The monopolist maximizes revenue by producing at the minimum rate in this interval. For $0 \leq Q \leq \bar{Q}$, the monopolist maximizes revenue by producing at the maximum rate in this interval. Hence the revenue maximizing production rate for the monopolist is

$$Q_{\ell} = \bar{Q} \quad .$$

It is clear that such a policy differs markedly from the socially optimal allocation, or the profit-maximizing monopoly strategy when $\alpha > 1$. In particular, for $\alpha < 1$ the policy is independent of the size of the monopoly resource stock. However, high prices encourage entry which erodes profits. The monopolist will choose a price strategy that balances current profits against increased future competition. The structure of dynamic limit pricing is developed, for example, in Gaskins,[7/] Phelps and Winter[8/] and Spence.[9/]

While this may seem somewhat far afield from the cartel-fringe problem that occupies us here, we shall show that in fact when $\alpha \leq 1$, the optimal cartel policy, assuming no uncertainty and perfect foresight, is described by a limit price determined by the supply response of the competitive fringe. The supply response depends on both the total stock of the resource owned by the fringe, and the output capacity of the fringe. Let us consider first the case in which the fringe is not constrained by production capacity.

We can determine the equilibrium allocation, assuming perfect foresight, as follows. If at any time,

$$\frac{\dot{P}}{P} < r \quad ,$$

the supply response of the fringe will force the price down to zero. Since the supply response of the cartel will insure that $\dot{P}/P \leq r$ (if $\dot{P}/P > r$ neither the cartel nor the fringe will produce and price will rise until $\dot{P}/P \leq r$), it follows that $\dot{P}/P = r$ as long as $S_f(t) > 0$ and $P(t) < \bar{P}$.

It is in the interest of the cartel to choose a pricing policy that encourages rapid depletion of the fringe stock, since the fringe stands in the way of monopoly profits. However, the cartel can only do this by choosing a low initial price, P_0, and this reduces its own profits in the short-run.

Since it doesn't matter to whom the fringe sells, we may suppose that the cartel abids by the pricing rule,

$$\frac{\dot{P}(t)}{P(t)} = r \quad \text{for } P(t) < \bar{P} \quad ,$$

and buys the entire fringe stock at $t = 0$ at price P_0. The cartel profits can then be described entirely by the choice of P_0. The cartel's maximum profits are

$$(9) \quad \Pi^{\ell} = \max_{P_0} \{-P_0 S_f + P_0 \int_0^T Q(P_0 e^{rt})dt + \frac{1}{r}\bar{P}e^{-rT}Q(\bar{P})[1 - \exp[\frac{-r(S_0 - \int_0^T Q(P_0 e^{rt})dt)}{Q(\bar{P})}]]\}$$

subject to

$$P_0 \leq \bar{P} \quad .$$

Let

$$S_0 \equiv S_\ell + S_f \quad .$$

Now

$$\bar{P}e^{-rt} = P_0 \quad ,$$

and

$$T = \frac{1}{r} \ln \frac{\bar{P}}{P_0} \quad .$$

Assume a constant elasticity demand function given by

(10) $$Q(P_0 e^{rt}) = \left(\frac{\varepsilon - 1}{P_0 e^{rt}}\right)^{\frac{1}{\varepsilon}} \quad ,$$

where

$$\varepsilon = \frac{1}{\alpha} > 1 \quad .$$

Now

$$\int_0^T Q(\cdot)\,dt = \int_{P_0}^{\bar{P}} Q(\cdot)\frac{dP}{\frac{dP}{dt}} = \int_{P_0}^{\bar{P}} \frac{Q(r)}{rP}\,dP = \int_{P_0}^{\bar{P}} \frac{1}{r}\left(\frac{\varepsilon - 1}{P}\right)^{\frac{1}{\varepsilon}} \frac{dP}{P} = \frac{(\varepsilon - 1)^{\frac{1}{\varepsilon}}}{r} \frac{1}{-\frac{1}{\varepsilon}} P^{-\frac{1}{\varepsilon}} \Big|_{P_0}^{\bar{P}}$$

Define

$$S_1 \equiv \frac{\varepsilon(\varepsilon - 1)^{\frac{1}{\varepsilon}}}{r}[P_0^{-\frac{1}{\varepsilon}} - \bar{P}^{-\frac{1}{\varepsilon}}]$$

$$[= 0 \text{ if } P_0 = \bar{P}] \quad .$$

Then

$$(11) \qquad \Pi^{\ell} = \max_{P_0} \{-P_0 S_f + P_0 \frac{\varepsilon(\varepsilon - 1)^{\frac{1}{\varepsilon}}}{r}(P_0^{-\frac{1}{\varepsilon}} - \bar{P}^{\frac{1}{\varepsilon}}) + \frac{1}{r} P_0 Q(\bar{P})[1 - \exp[\frac{-r}{Q(\bar{P})}(\frac{S_0 - \frac{\varepsilon(\varepsilon - 1)^{\frac{1}{\varepsilon}}}{r}(P_0^{-\frac{1}{\varepsilon}} - \bar{P}^{-\frac{1}{\varepsilon}})}{S_0 - S_1(P_0)})]]\} \quad .$$

Necessary and sufficient conditions for a maximum are

$$(12) \qquad P_0^{-\frac{1}{\varepsilon}} \varepsilon(\varepsilon - 1)^{\frac{1}{\varepsilon}} [(\varepsilon - 1) + e^{-r\frac{(S_0 - S_1(P_0))}{Q(\bar{P})}}] + \frac{1}{r} Q(\bar{P})[1 - e^{-r(\frac{S_0 - S_1(P_0)}{Q(\bar{P})})}] = rS_f + \frac{\varepsilon(\varepsilon - 1)^{\frac{1}{\varepsilon}}}{r} \bar{P}^{-\frac{1}{\varepsilon}} \quad .$$

Now

$$(13) \qquad Q(\bar{P}) = (\frac{\varepsilon - 1}{\bar{P}})^{\frac{1}{\varepsilon}} = (\varepsilon - 1)^{\frac{1}{\varepsilon}} \bar{P}^{-\frac{1}{\varepsilon}} \quad .$$

Substituting (13) in (12) and rearranging terms gives

$$(14)\quad P_0^{\ell} = \left\{ \frac{\dfrac{rS_f}{(\varepsilon - 1)^{\frac{1}{\varepsilon}}} + \dfrac{1}{\bar{P}^{\frac{1}{\varepsilon}}}\left[(\varepsilon - 1) + e^{-r\frac{(S_0 - S_1(P_0))}{Q(\bar{P})}}\right]}{(\varepsilon - 1) + e^{-r\frac{(S_0 - S_1(P_0))}{Q(\bar{P})}}} \right\}^{-\varepsilon} .$$

Suppose the length of time required to exhaust $S_0 - S_1$ is sufficiently large that

$$(\varepsilon - 1) >> e^{-r\frac{(S_0 - S_1(P_0))}{Q(\bar{P})}}$$

Then a reasonable approximation to (14) is

$$(15)\qquad P_0^{\ell} = (\varepsilon - 1)\left[\frac{rS_f}{\varepsilon - 1} + \left(\frac{\varepsilon - 1}{\bar{P}}\right)^{\frac{1}{\varepsilon}}\right]^{-\varepsilon} .$$

Compare this to the efficient P_0^* for the fringe stock S_f:

$$(16)\qquad P_0^*(S_f) = (\varepsilon - 1)\left[\frac{rS_f}{\varepsilon} + \left(\frac{\varepsilon - 1}{\bar{P}}\right)^{\frac{1}{\varepsilon}}\right]^{-\varepsilon} .$$

It follows that

$$P_0^{\ell} < P_0^*(S_f) \quad .$$

Also, it must be true that

$$P_0^{\ell} > P_0^*(S_0)$$

since the cartel can choose to supply the resource at the efficient competitive price. For this case, the cartel policed by a competitive fringe with unlimited production capacity chooses a price path $P(t)$ described as follows:

(i) The rate of change of $P(t)$, $\dot{P}/P$ is equal to r, the efficient rate of return on the resource stock. However,

(ii) the magnitude of the price, $P(t)$, is bounded below by the efficient price for the total endowment, S_0, and bounded above by the efficient price for the stock limited to the amount owned by the fringe, S_f. This is illustrated in figure 16.

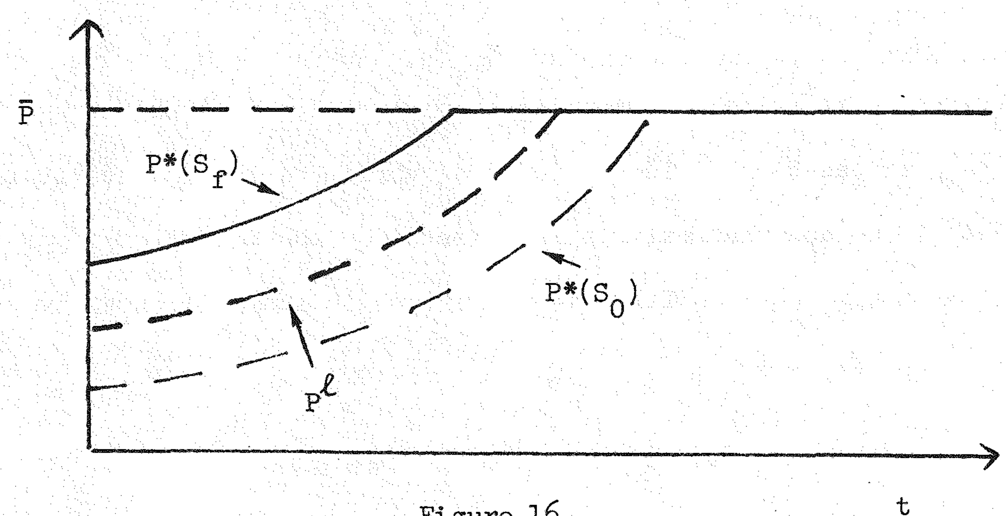


Figure 16

Note that the price leader allocates efficiently for an <u>equivalent</u> stock of $(\varepsilon/(\varepsilon - 1))S_f$ with backstop price $\bar{P}$. Because $P_0^{\ell} < P_0^*(S_f)$, and $\dot{P}/P = r$ for $P < \bar{P}$, it follows that

$$\int_0^T Q(P(t)dt > S_f \quad ,$$

where T is defined by $P(T) = \bar{P}$. In other words, the cartel produces some of its own stock during $t \in (0,T)$, in addition to all of the fringe stock. This is clear from

$$S_{eq} = \frac{\varepsilon}{\varepsilon - 1} S_f > S_f \quad \text{for} \quad \varepsilon < \infty \quad .$$

The amount in excess of S_f is just

$$\left(\frac{\varepsilon}{\varepsilon - 1} - 1\right) S_f = \frac{1}{\varepsilon - 1} S_f \quad .$$

Adelman[10/] has calculated that the OPEC cartel enjoys a considerable cost advantage relative to competing suppliers of petroleum, as well as claiming over one-half of identified free world reserves. Thus it is of interest to see how the analysis is affected by the inclusion of production costs. Assume production is constant returns to scale if capacity constraints are not binding. Let

C_f = unit production costs of the fringe.

C_ℓ = unit production cost of the cartel.

We assume $C_f > C_\ell$.

If fringe capacity is unlimited, then, by the previous argument, the cartel must set prices so that

$$\frac{\dot{P}_t}{P_t - C_f} = r \quad \text{for} \quad P(t) < \bar{P} \quad .$$

The cartel's pricing strategy is independent of C_ℓ, since the profit maximizing price is $\bar{P}$ for $C_\ell < \bar{P}$. The derivation of the optimal cartel pricing strategy proceeds with only minor changes, and the equilibrium extractions path is as shown in figure 17.

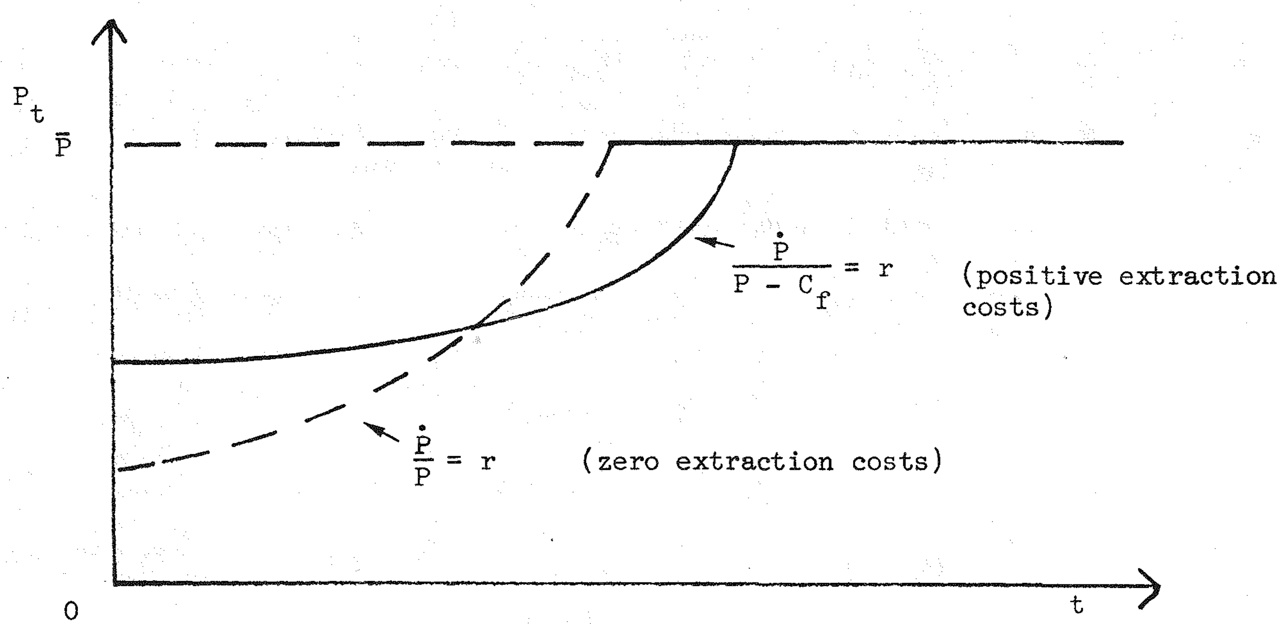


Figure 17

It should be stressed that the resource allocation determined in this section would only obtain in practice if contracts for all future dates could be made and enforced. For example, if the cartel did buy out the fringe, it would, if not policed, depart from the agreed upon extraction policy and set price equal to $\bar{P}$ to maximize profits. If the fringe suspected that this would be the result, competitors would not sell the resource to the cartel in the first place. However, the cartel can provide its own guarantee, namely by producing its own stock at the equilibrium price. The fringe can then schedule production to exhaust stocks when price equals $\bar{P}$. If the cost of extraction from the fringe stock exceeds that of the cartel, production in the absence of futures markets would be efficient since the lower cost deposits would be extracted first.

This result depends on perfect information about the remaining supplies in the fringe. Given the uncertainty in world supplies and the policy decision of producers it is unlikely that the delicate equilibria described in this section would obtain. (Nonetheless, the solutions emphasize the incentives bearing on suppliers of the exhaustible resource.) Suppose instead that the cartel chose a constant price and the fringe responds by producing at capacity, at least until it becomes evident that the total supply of the fringe will be exhausted in the future and the price will rise suddenly to $\bar{P}$. We examine this outcome in the next section.

$\alpha \leq 1$; fringe capacity limited

The cartel's profit-maximizing strategy takes account of the supply response of the fringe. If the cartel sets a price trajectory with

$$\frac{\dot{P}}{P} < r \quad ,$$

the fringe will respond by producing at capacity, k_f. Assume the size of the cartel stock is sufficiently large so that the date of exhaustion, T, is far enough away that $rT >> 0$. In than case, the cartel maximizes profits by choosing q_ℓ to

(17) $$\max_{q_\ell} P(q_\ell + k_f)q_\ell \quad ,$$

subject to the constraint that

$$P(q_\ell + k_f) \leq \bar{P} \quad .$$

The cartel could of course elect to set the price trajectory described in the previous section.[11/] Comparison of the strategies will be postponed until the end of this section. We will show that the fringe capacity constraint enables the cartel to increase its profits. However, if the fringe producers have rational expectations, and if the fringe stock is exhausted before the cartel stock, the fringe supply response eventually will force the cartel to conform to the price schedule described in the preceding section.

Maximization of (17) requires

$$P(q_\ell + k_f)[1 - \frac{q_\ell}{q_\ell + k_f}\frac{1}{\alpha} + \lambda\frac{1}{q_\ell + k_f}\frac{1}{\alpha}] = 0$$

or

$$\hat{q}_\ell = \frac{\lambda}{1 - \alpha} + \frac{\alpha}{1 - \alpha}k_f \quad ,$$

where λ is the shadow price to the cartel of the fringe capacity constraint.

The solution for this policy choice looks as shown in figure 18.

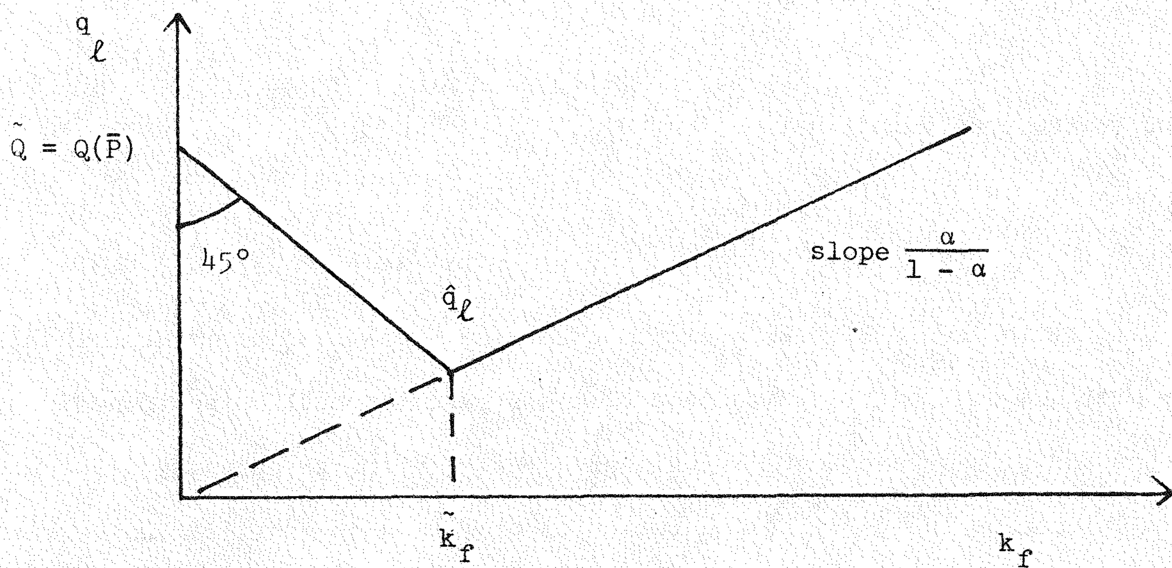


Figure 18

The non-monotonicity of the q_ℓ is the result of the constraint that

$$P(q_\ell + k_f) \leq \bar{P} \quad .$$

In the absence of the constraint the profit-maximizing q_ℓ would be proportional to k_f, with slope

$$\frac{\alpha}{1 - \alpha}$$

However, for $k_f < k_f$, the cartel is limited by the backstop price $\bar{P}$. For this strategy, given $\bar{P}$, the resource price dependence on k_f is as shown in figure 19.

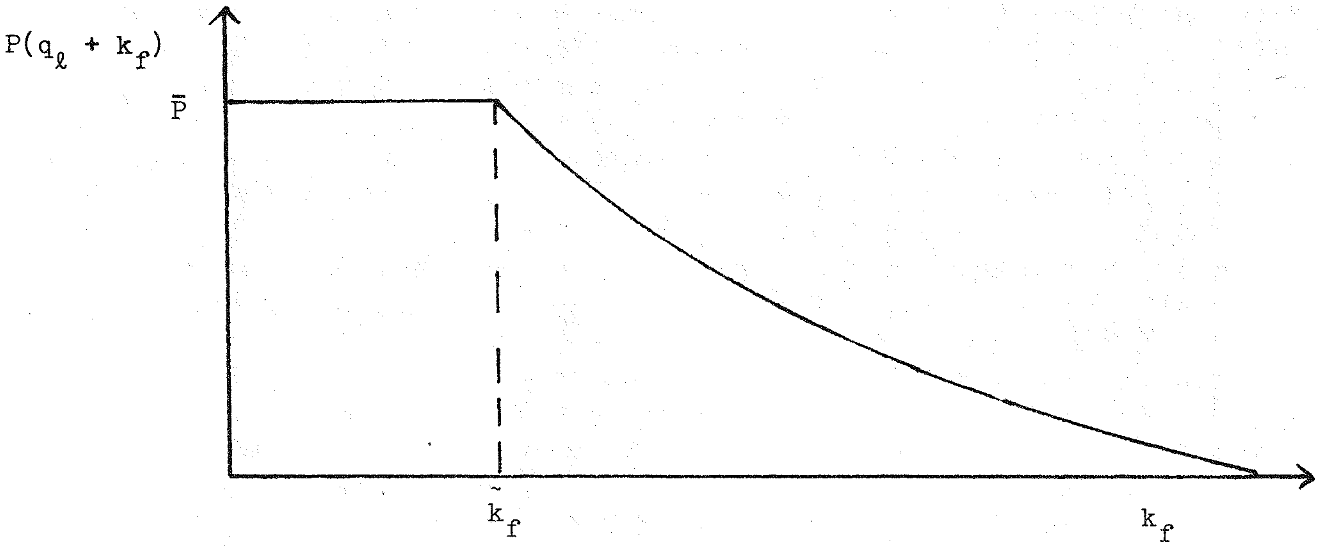


Figure 19

The dependence of cartel revenues on k_f is

$$R_\ell(k_f) = \bar{P}[Q|\bar{P}) - k_f] \quad \text{for} \quad k_f \in [0,\tilde{k}_f] \quad ,$$

and

$$R_\ell(k_f) = \frac{\alpha}{1-\alpha} P(\frac{k_f}{1-\alpha})k_f \quad \text{for} \quad k_f > \tilde{k}_f \quad .$$

This is illustrated in figure 20.

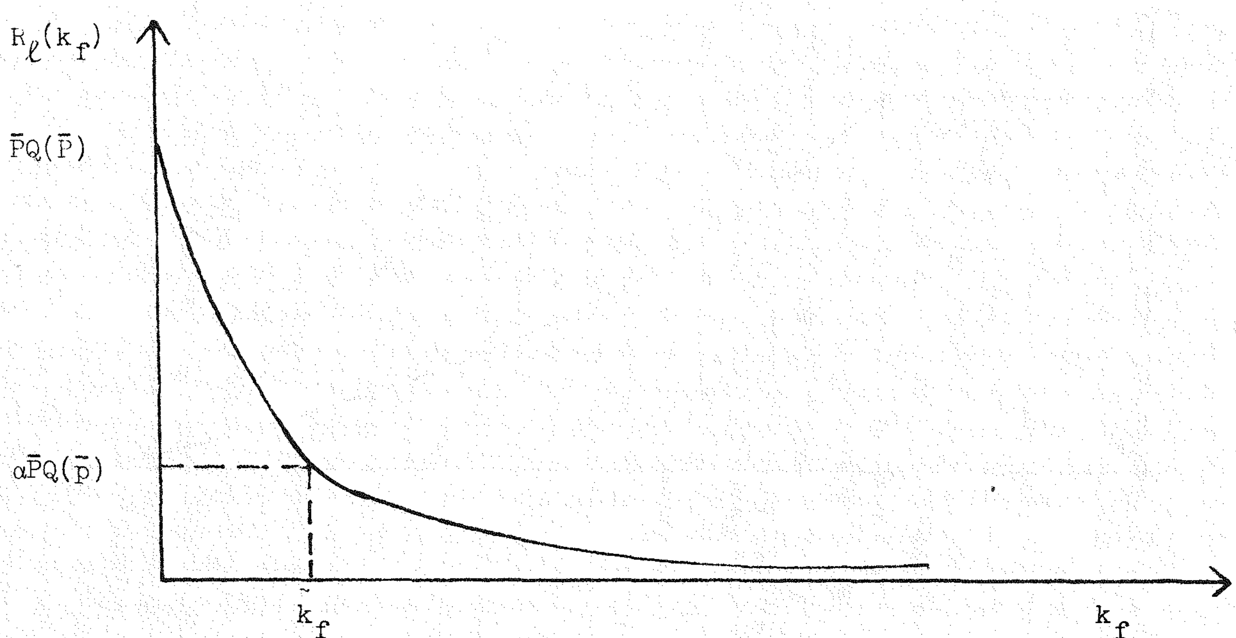


Figure 20

The cartel could choose either a constant price, or the price policy discussed in section 7.2.3 Let us suppose that producers comprising the fringe were myopic in their supply responses to prices set by the cartel. In that case, the cartel could choose between the following revenue streams.

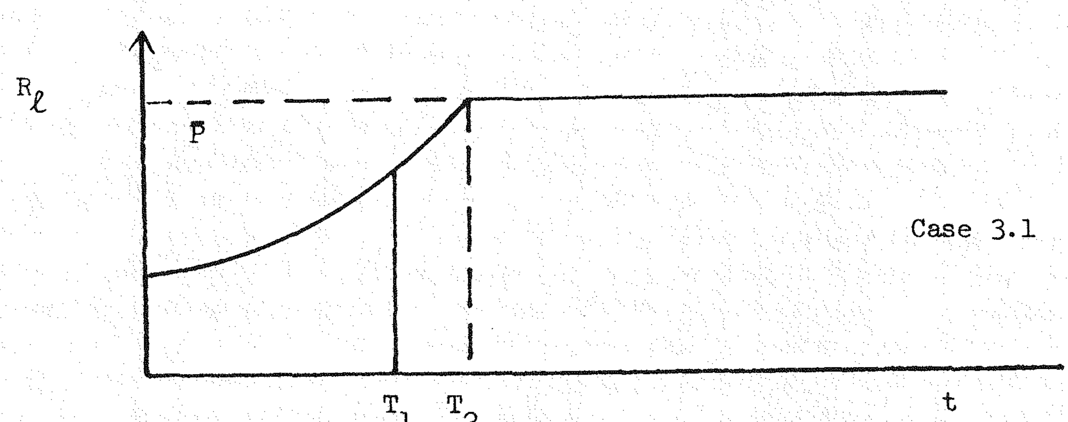


Figure 21

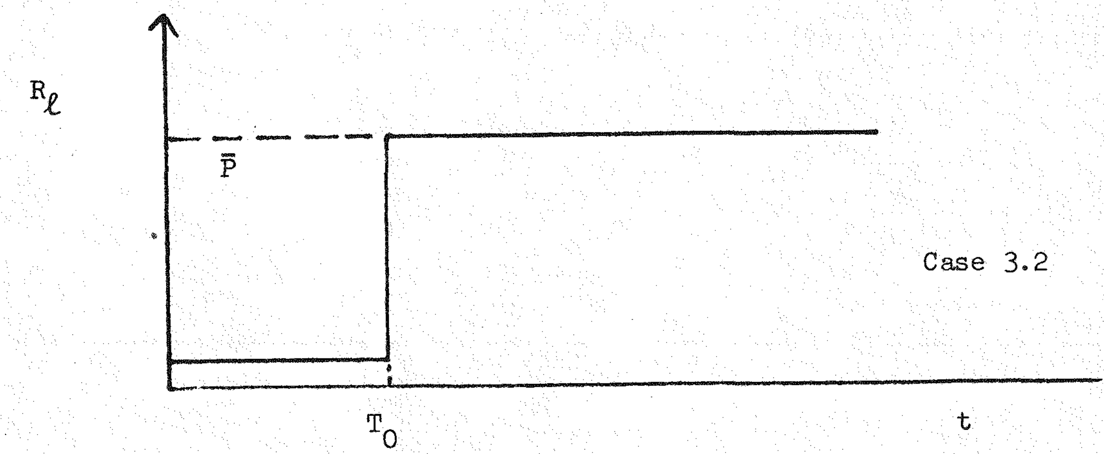


Figure 22

If the cartel stock were large relative to S_f, the revenue stream shown above in case 3.2 would dominate that of case 3.1. The reason is that in case 3.2, the cartel exhausts the fringe as soon as possible. In case 3.1, the cartel had to buy out the fringe, and does not make up for this outlay until time T_1. (The purchase need not actually occur, however, the amount of the purchase is the reduction in cartel profits that result from production by the fringe.) The cartel actually does better in case 3.2, provided it has enough of the resource to spare in the flurry of production that occurs before time T_0.

However, case 3.2 is not an intertemporal equilibrium allocation if competitors are not myopic in their supply decisions. Competitive producers or the resource who anticipate the price rise at time T_0 would hold back supply, and the market price would rise. Let $P_1(t)$ be the price trajectory for case 3.1. Suppose that the revenue-maximizing cartel price when the fringe produces at maximum capacity exceeds $P_1(t)$. Then there would be a period of constant prices as the fringe stock is depleted, followed by rising prices until the fringe stock is exhausted at $P(T) = \bar{P}$ (see figure 23).

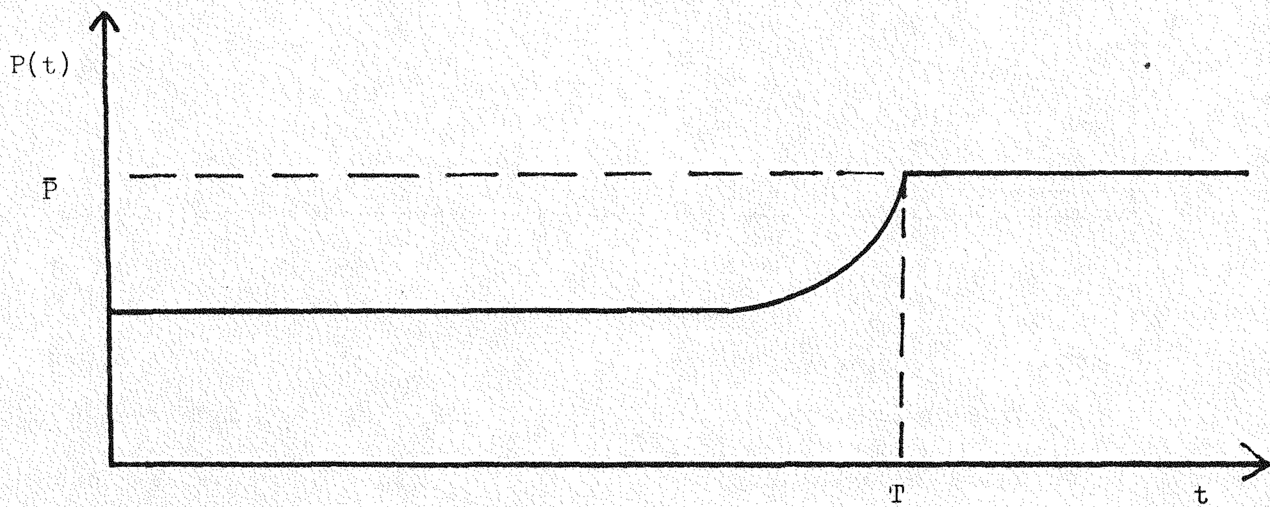


Figure 23

Whether or not the resource price exhibits a rising trend in the immediate future depends in this formulation on the capacity limitation of competitive suppliers of the resource. However, in all cases there is a tendency for prices to rise eventually toward the cost of a substitute source of supply. The costs of substitutes ultimately determine the monopoly power of the resource cartel. We have assumed in the preceding analysis that a substitute was available at a constant marginal cost. In the next section, we examine the returns from development of a substitute for an exhaustible resource when there are fixed costs associated with the substitute production technology.

Section 3: Footnotes

1/ Elsewhere in this report we use α for the magnitude of the demand elasticity. That is, $\eta = -\alpha$.

2/ Extraction costs change this result as discussed in the first interim report, still, the important result is that the distortionary effect of the cartel is limited when demand is elastic $(\alpha > 1)$.

3/ See, for example, Berndt, E. and D. Wood, "Technology, Prices and the Derived Demand for Energy," University of British Columbia Discussion Paper 74-09, May 1974, and Fuss M.. R. Hyndman, and L. Waverman, "Residential, Commercial and Industrial Demand for Energy in Canada," Stanford University mimeo, 1976.

4/ Note: we assumed $\partial P/\partial t > 0$.

5/ We shall refer interchangeably to a cartel, monopoly or price-leader to describe the sector sharing the market with the competitive fringe.

6/ In section 5.2 of this report we show that for some circumstances, the profit maximizing monopoly strategy is to randomize prices.

7/ Gaskins, D., "Dynamic Limit Pricing," Journal of Economic Theory, 1971.

8/ Phelps, J., and S. Winter [1971], "Limit Pricing under Atomistic Competition," in Phelps and Winter, (eds) The Microeconomic Foundations of Macroeconomics.

9/ Spence, M., "Entry, Capacity, and Oligopolistic Pricing," IMSSS Technical Report #131, Stanford University, 1974.

10/ Adelman, M., The World Petroleum Market, Johns Hopkins Press, 1971.

11/ The only other possibility is $\dot{P}/P > r$, but in that case both the cartel and the fringe will hold back production.

8. Increasing Returns in the Development of Substitutes

When a substitute for an exhaustible resource is available and can be produced at constant marginal cost, $\bar{P}$, we know from the discussion in the first interim report that the efficient resource price will rise to meet $\bar{P}$ when the resource is exhausted. Furthermore, we have shown that the efficient allocation is sustained in a perfectly competitive economy with rational expectations.

Previous discussion has neglected the fixed costs associated with the introduction of a substitute. In section 8.1, we determine the socially optimal management of an economy with an exhaustible resource stock (with zero extraction cost) and a substitute that can be produced at constant marginal cost $\bar{P}$, but requires a fixed cost, A, for its introduction. In section 8.2 we extend the analysis to consider variable development costs that depend on the anticipated date at which the substitute will be introduced.

For an efficient allocation, the substitute must be priced at its marginal production cost, $\bar{P}$. This means that if the substitute is supplied competitively, there are no profits to cover the cost of developing the substitute technology. If development is socially managed, the government finances the cost through taxation. Patent rights are relied upon to provide incentives for decentralized research and development. The patent grants monopoly power (for a limited time) to the firm that introduces the invention. In section 8.3, we analyze the profit-maximizing date at which a substitute would be introduced by a monopolist on both the exhaustible resource and the substitute source supply. The case of competition for an infinite patent (i.e. monopoly) on the substitute is the subject of section 8.4.

In section 8.4 we join the problem of decentralized development of a substitute source of energy with the analysis of cartel pricing policy. The cartel limit price depends on the introduction of substitutes. However, even with an infinite patent, there may be no incentive for decentralized development because the cartel can price just below the production cost of the substitute.

8.1 Socially optimal introduction of the substitute

Define:

- A: fixed development cost of the substitute for the exhaustible resource.
- $V(S_0,T)$: social value of the exhaustible stock, S_0, depleted over a time horizon, T.
- $V(\bar{P})$: social value of an inexhaustible substitute with production cost $\bar{P}$.

Let

(1) $$W^S(S_0,T) = V(S_0,T) + e^{-rT}[V(\bar{P}) - A] \quad .$$

$W^S(S_0,T)$ is the social return from introduction of the substitute at date T.

Define $\bar{h}$ by

(2) $$V(\bar{h}) = V(\bar{P}) - A \quad .$$

By construction, $\bar{h}$ is the unit cost of a constant returns to scale (CRS) technology that yields the same present social value as the technology with development cost A and unit production cost $\bar{P}$.

Assume the demand price of the resource is given by

(3) $$P(Q) = Q^{-\varepsilon} ,$$

where

$$\varepsilon = \frac{1}{\alpha} .$$

Then for $\alpha > 1$,

(4) $$\bar{h} = \frac{\bar{P}}{\left[1 - \frac{r}{\varepsilon}\left(\frac{\bar{P}}{1 - \varepsilon}\right)^{\frac{1-\varepsilon}{\varepsilon}} A\right]^{\frac{\varepsilon}{1-\varepsilon}}} ,$$

and for $\alpha < 1$

(5) $$\bar{h} = \bar{P}\left[1 + \frac{r}{\varepsilon}\left(\frac{\varepsilon - 1}{\bar{P}}\right)^{\frac{\varepsilon-1}{\varepsilon}} A\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

The resource price trajectory that supports the optimal allocation is shown in figure 1

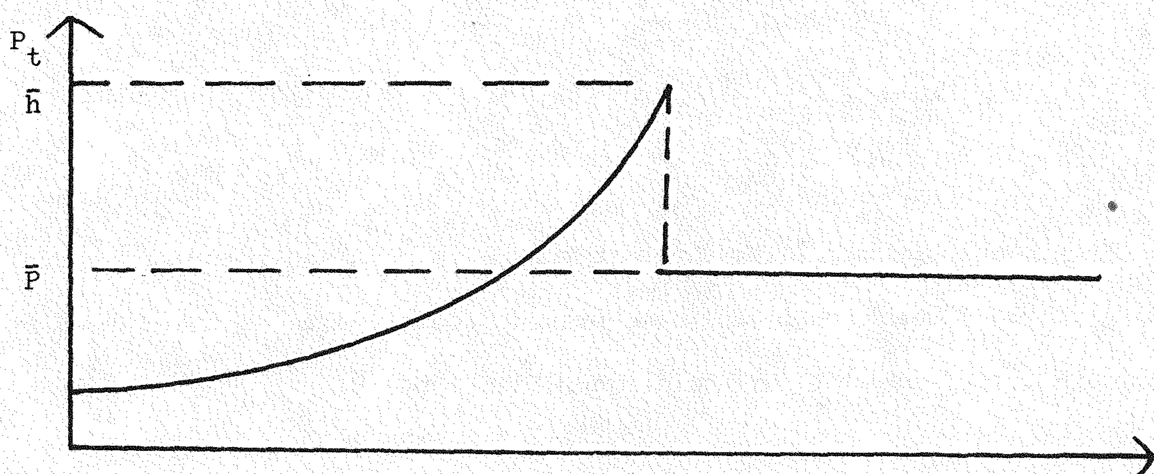


Figure 1

Note that there is an interval of time during which the resource price exceeds the marginal cost of the substitute, and at time T there is a sudden change in the price corresponding to introduction of the substitute. Suppose, for example, that the current price of crude oil exceeds the marginal variable cost of shale oil. Given the costs of introducing the shale oil technology, the analysis here suggests that the price differential is not sufficient cause to argue that the introduction of shale oil should be accelerated.

8.2 Endogenous Availability of Substitute: The Socially-Managed Economy

In this section, we consider the somewhat more general case in which the time at which a substitute may be introduced is a function of the level of investment. The optimal investment plan when both the entire exhaustible resource stock and the substitute technology are socially managed is determined. It is shown that the optimal level of investment may be a discontinuous function of time.

The socially-managed case is a benchmark for comparison of investment in substitute development when both the resource and the substitute are controlled by a monopolist (section 8.3), and when the resource is owned competitively and there is competition for a patent on the substitute technology (section 8.4).

The present social value, gross of development costs, of a resource stock S_0 and a substitute introduced at time T is $W^s(S_0,T)$ defined by

(6) $$W^s(S_0,T) = V(S_0,T) + e^{-rT}V(\bar{P}) \quad .$$

For the assumed demand conditions,

$$(7)\qquad W^s(S_0,T) = \left(\frac{\varepsilon}{r}\right)^{\varepsilon} S_0^{1-\varepsilon}\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{\varepsilon} + e^{-rT}\frac{\varepsilon}{r}\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}} .$$

Investment in the development of the substitute is determined to

$$(8)\qquad \max_x \{W^s(S_0,T(x)) - x\} .$$

$$\text{s.t.}\quad x \geq 0 ,$$

where x is the level of investment in the development of the substitute.[1/]

First-order necessary conditions for a maximum are

$$(9)\qquad \varepsilon\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{\varepsilon-1}\left(\frac{r}{\varepsilon}\right)e^{-\frac{rT}{\varepsilon}}T'\left(\frac{\varepsilon}{r}\right)^{\varepsilon} S_0^{1-\varepsilon} - re^{-rT}T'\frac{\varepsilon}{r}\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}} = 1 - \lambda ,$$

where λ is the lagrange multiplier for the constraint $x \geq 0$. Therefore,

$$(10)\qquad \frac{S_0^{1-\varepsilon}e^{-\left(\frac{1-\varepsilon}{\varepsilon}\right)rT}}{\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}} = \frac{(1-\lambda)}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\frac{e^{rT}}{T'} + \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}} .$$

Assume $T(x)$ looks as shown in figure 2,

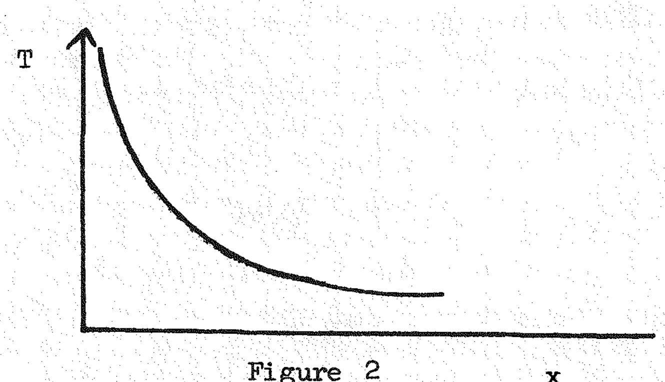


Figure 2

so that $T(0) \to \infty$. Then for $S_0 \leq \bar{S}_0$, $x \geq 0$ (i.e., $T(x) < \infty$) and $\lambda = 0$. Therefore, for $S_0 \leq \bar{S}_0$

$$(11) \qquad S_0^{1-\varepsilon} = \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon} \cdot \left\{\frac{e^{\frac{rT}{\varepsilon}}}{T'} + \varepsilon\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}} e^{\left(\frac{1-\varepsilon}{\varepsilon}\right)rT}\right\} .$$

Consider the right hand side (RHS) of

$$\text{as } x \to 0\ [T(x) \to \infty], \text{RHS } (11) \to \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon} \frac{e^{\frac{rT}{\varepsilon}}}{T'} \to -\infty$$

$$\text{as } x \to \infty\ [T(x) \to 0], \text{RHS } (11) \to \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\left[\frac{\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}}{T'} + \varepsilon\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}}\left(1-e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}\right]$$

$\to -\infty$ if $T(x)$ is constant elasticity.

If $T(x)$ is constant elasticity, it can be shown that

(a) either there is no x for which the RHS of (11) $= 0$ ($\Rightarrow x^* = 0$ for all $S_0 \geq 0$), or

(b) there are two values of x for which the RHS of (11) $= 0$.

The RHS of (11) looks like (for case (b)):

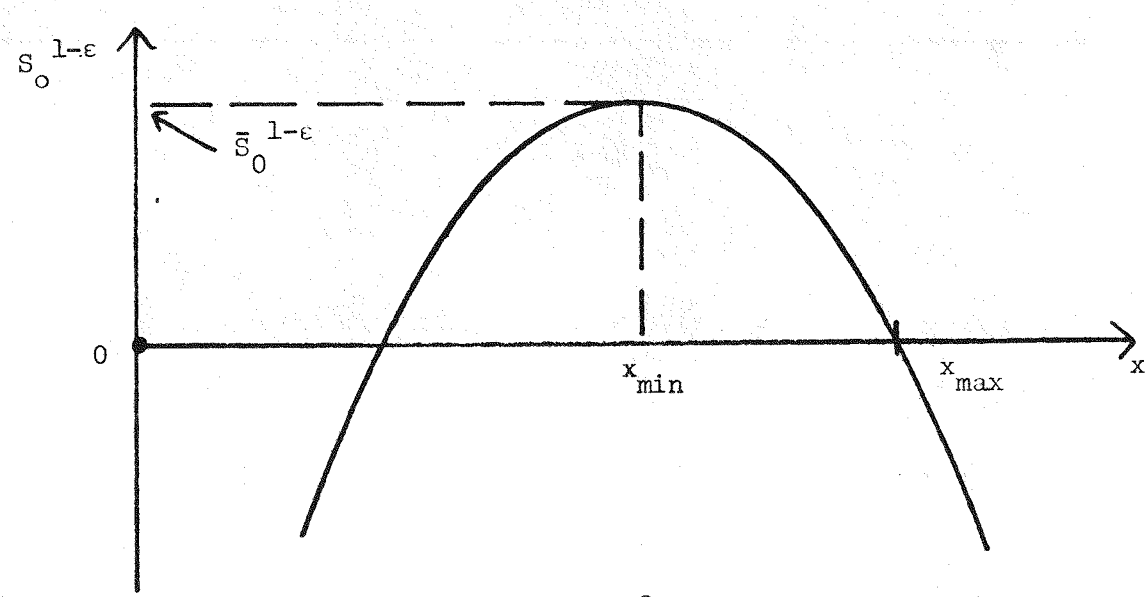


Figure 3

For $S_0 > \bar{\bar{S}}_0^{1-\varepsilon}$, $x^* \equiv 0$. (The constraint $x \geq 0$ is binding.)

We can see why this is so by plotting the family of curves $W^S(S_0,T(x))$ as a function of x for different values of S_0 (figure 4).

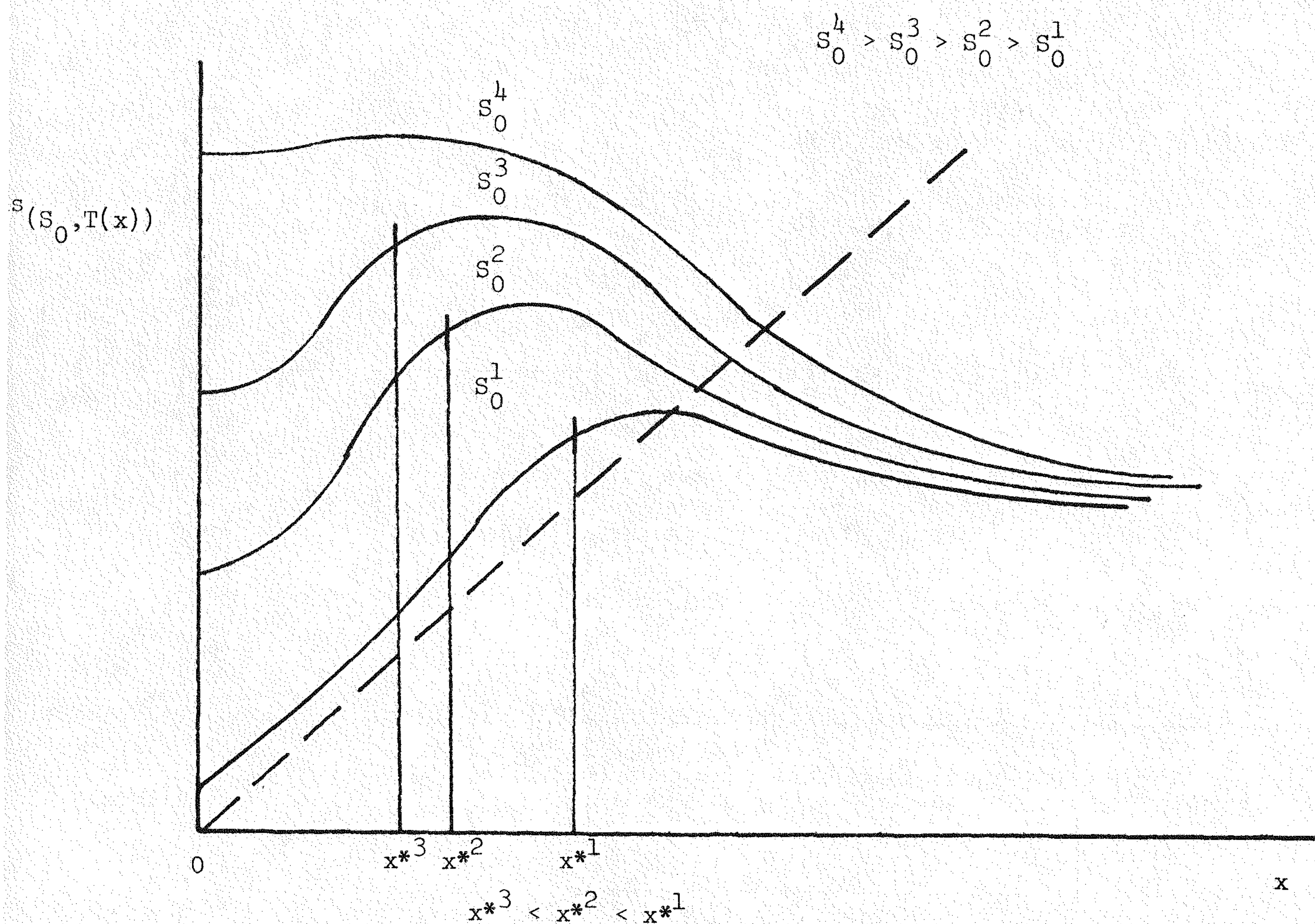


Figure 4

It is the higher value of x (in the two solutions to (11)) that maximizes net surplus.

For $\bar{S}_0$ there is only one value of x that is a solution to (11). For $S_0 > \bar{S}_0$, there is no value (e.g., S_0^4 above). I.e., the marginal benefit of investment is less than one for $S_0 > \bar{S}_0$, and the marginal benefit of investment equals one for $S_0 = \bar{S}_0$. At this point, $x^*(\bar{S}_0) > 0$ because of the non-convexity in the benefit function (terms like $T'e^{-rT}$).[2]

The optimal investment in the substitute as a function of the remaining resource stock, S_0, is as shown in figure 5.

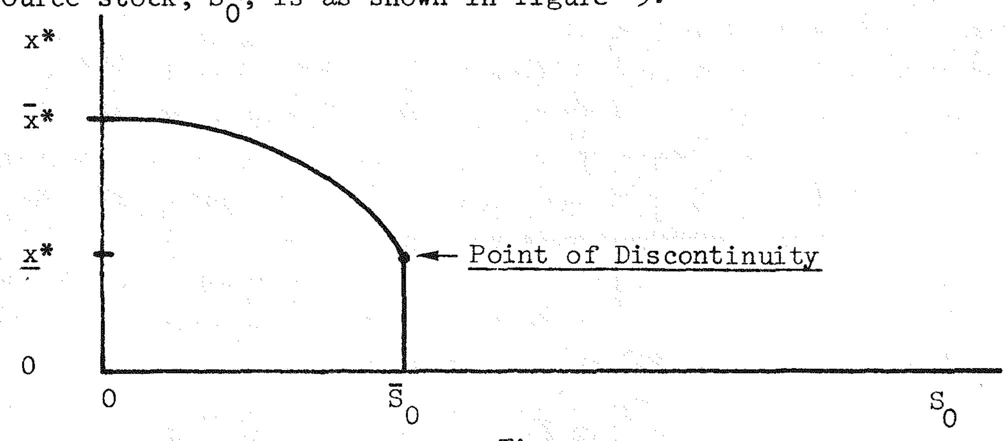


Figure 5

Also, prices must be discontinuous -- because a continuous price implies zero investment in the arrival time of the substitute.

8.3 Monopoly vs. Social Optimum

The monopolist's net revenue from the stock S_0 and the substitute introduced at time T is

$$(12)\quad W^m(S_0,T(x)) = (1 - \varepsilon)\left(\frac{\varepsilon}{r}\right) S_0^{1-\varepsilon}\left(1 - e^{-\frac{rT}{\varepsilon}}\right)^{\varepsilon} + \varepsilon e^{-rT}\left(\frac{1 - \varepsilon}{r}\right)^{\frac{1}{\varepsilon}}\left(\frac{1 - \varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}}$$

The difference between marginal net monopoly revenue and marginal social value of investment in the substitute is

$$(13)\quad W_x^m - W_x^s = \frac{dT}{dx}\left[\left\{e^{-rT}\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}} \varepsilon\left[1-(1-\varepsilon)^{\frac{1}{\varepsilon}}\right]\right\} - \left\{\frac{S_0^{1-\varepsilon} r\varepsilon\left(\frac{\varepsilon}{r}\right)^{\varepsilon} e^{-\frac{rT}{\varepsilon}}}{\left(1-e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}}\right\}\right]$$

Since $dT/dx < 0$, $W_x^m - W_x^s \gtrless 0 \Rightarrow$

$$(14)\quad \frac{S_0^{1-\varepsilon} e^{\left(\frac{1-\varepsilon}{\varepsilon}\right) rT}}{\left(1-e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}} \gtrless \frac{\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}}\left[1-(1-\varepsilon)^{\frac{1}{\varepsilon}}\right]}{r\left(\frac{\varepsilon}{r}\right)^{\varepsilon}}$$

From (11) (first order necessary conditon for x*), we have

$$(15)\quad \frac{S_0^{1-\varepsilon} e^{-\left(\frac{1-\varepsilon}{\varepsilon}\right) rT}}{\left(1-e^{-\frac{rT}{\varepsilon}}\right)^{1-\varepsilon}} = \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\left[\frac{e^{rT}}{T'} + \varepsilon\left(\left(\frac{1-\varepsilon}{\bar{p}}\right)\right)^{\frac{1-\varepsilon}{\varepsilon}}\right]$$

Substituting (15) in (14) gives

$$(16)\quad \frac{1}{\varepsilon}\left(\frac{\varepsilon}{r}\right)^{1-\varepsilon}\left[\frac{e^{rT}}{T'} + \varepsilon\left(\left(\frac{1-\varepsilon}{\bar{p}}\right)\right)^{\frac{1-\varepsilon}{\varepsilon}}\right] \gtrless \frac{\left(\frac{1-\varepsilon}{\bar{p}}\right)^{\frac{1-\varepsilon}{\varepsilon}} \varepsilon\left[1-(1-\varepsilon)^{\frac{1}{\varepsilon}}\right]}{r\left(\frac{\varepsilon}{r}\right)^{\varepsilon}},$$

or

$$(17) \qquad \frac{e^{rT(x)}}{T'(x)} \gtrless \left(\frac{1-\varepsilon}{\varepsilon}\right)\left[\frac{1-\varepsilon}{\bar{P}}\right]^{\frac{1-\varepsilon}{\varepsilon}}\left[1-(1-\varepsilon)^{\frac{1-\varepsilon}{\varepsilon}}\right]$$

Since the right hand side of (17) is strictly positive, and the left hand side is strictly negative, it follows that

$$W_x^m < W_x^s \ .$$

For all values of S_0, monopolist spends less on research for introduction of substitute than is socially optimal.

(8.4) Competitive Development of Substitute: Socially-Managed Resource

Suppose that the results of investment in a substitute source of supply is a known function of the investment level, and all firms are identical. If there were N firms engaged in development, an equilibrium would have each firm maximizing expected profits, which must be 1/N of total profits. However, this is not a stable equilibrium if expected profits are positive, since an increment of investment by any one firm would capture total profits. The only equilibrium is one firm engaged in development, at a level so that profits are zero to forestall entry. The private incentives for development of the substitute depend on the future price of the resource. We have assumed the resource has zero extraction cost. Hence the resource can always be priced to undercut the price of the substitute. In that case, we may suppose that the innovating firm announces the arrival date of the substitute, $T(x_c)$, and resource production is scheduled so that the stock is exhausted at $T(x_c)$.

Given the announced arrival date of the substitute, $T(x_c)$, the resource would be produced efficiently if owned by competitors as well as if it were socially managed. However, the results of this section may be changed dramatically if the resource stock is owned by a monopolist, as we shall see in the next section.

It is assumed that an infinite patent is awarded to the firm that develops the substitute (inclusion of finite patent rights is straightforeward). In this case, the value of the development program is

$$(18) \quad e^{-rT(x_c)}\frac{\varepsilon}{r}(1-\varepsilon)^{\frac{1}{\varepsilon}}\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}} \quad \text{for} \quad \varepsilon < 1 \quad .$$

For zero profits, the firm will equate investment cost and revenues, so that

$$(19) \quad x_c = e^{-rT(x_c)}\frac{\varepsilon}{r}(1-\varepsilon)^{\frac{1}{\varepsilon}}\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}}$$

Let us compare the allocation of investment in development, x_c, with the optimal allocation, x^*. For x^*,

$$(20) \quad \max_{x^*} W^S = \max_{x^*} \{[(\frac{\varepsilon}{r})^{\varepsilon} S_0^{1-\varepsilon}(1 - e^{-\frac{rT}{\varepsilon}})\varepsilon + e^{-rT}\frac{\varepsilon}{r}(\frac{1-\varepsilon}{\bar{P}})^{\frac{1-\varepsilon}{\varepsilon}}] - x\} \quad .$$

For x_c,

$$(21) \quad x_c = e^{-rT}\frac{\varepsilon}{r}(1-\varepsilon)^{\frac{1}{\varepsilon}}\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}}$$

provided $S(T(x_c)) = 0$.

As long as the innovating firm does not have to wait to introduce the substitute, the firm will ignore the effect of the substitute on the value of the exhaustible stock. Since earlier introduction of the substitute may lower the value of the stock, this implies that the competitive allocation to development may exceed the optimal level.

Define $\tilde{x}$ by

$$(22) \qquad \max_x \left[e^{-rT}\frac{\varepsilon}{r}\left(\frac{1-\varepsilon}{\bar{\bar{P}}}\right)^{\frac{1-\varepsilon}{\varepsilon}} - x\right] = \max_x W(S_0 = 0)$$

Then

$$\tilde{x} \geq x^*$$

Since

$$V(\bar{P}) = \frac{\varepsilon}{r}\left(\frac{1-\varepsilon}{\bar{P}}\right)^{\frac{1-\varepsilon}{\varepsilon}}$$

we have

$$(23) \qquad -r\frac{dT(\tilde{x})}{dx}e^{-rT(\tilde{x})}V(\bar{P}) = 1 \quad ,$$

which determines the optimal investment when the resource stock is zero, and

$$(24) \qquad e^{-rT(x_c)}(1-\varepsilon)^{\frac{1}{\varepsilon}}V(\bar{P}) = x_c$$

for the competitive equilibrium.

To see the relationship more clearly, rewrite (22) and (24) as:

For $\tilde{x}$

$$\max_{\tilde{x}} \left[e^{-rT(\tilde{x})} - \frac{\tilde{x}}{V(\bar{P})}\right] V(\bar{P}) \quad .$$

For x_c,

$$(1 - \varepsilon)^{\frac{1}{\varepsilon}} e^{-rT(x_c)} = \frac{x_c}{V(\bar{P})} \quad .$$

The relative magnitudes of $\tilde{x}$ and x_c depend on $V(\bar{P})$, ε, and the correspondence $T(x)$. This is illustrated in figure 6.

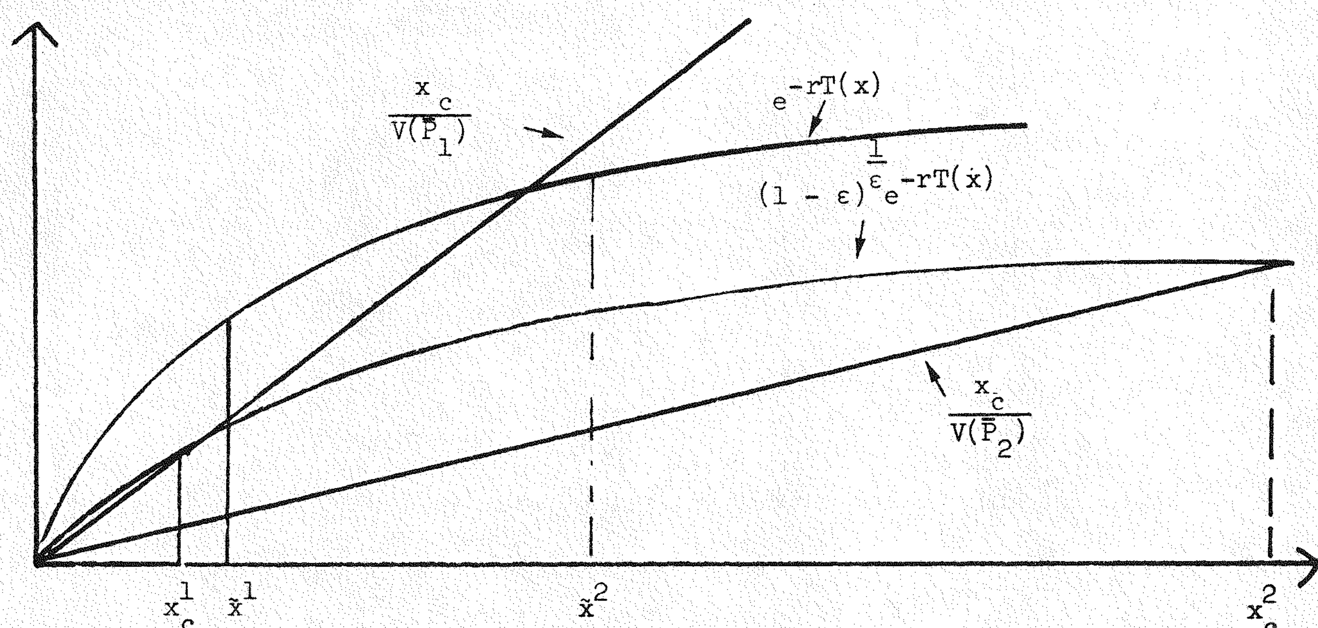


Figure 6

In figure 6,

$$x_c^1 < \tilde{x}^1 \quad ,$$

while

$$x_c^2 > \tilde{x}^2 \quad .$$

Consider the special case

$$e^{-rT(x)} = \frac{x}{x + 1}$$

$$x_c = (1 - \varepsilon)^{\frac{1}{\varepsilon}}(V(\bar{P}) - 1$$

Now

$$\frac{dW(S_0 = 0)}{dx} = \frac{d}{dx}[\, e^{-rT(x)}] \cdot V(\bar{P}) - 1 \ ,$$

or

$$\frac{dW}{dx} \mid x_c = \frac{1}{(1 - \varepsilon)^{\frac{2}{\varepsilon}} V(\bar{P})} - 1 \quad .$$

If

$$\frac{dW}{dx}\Big|_{x_c} > 0$$

then $\tilde{x} > x_c$, while

$$\frac{dW}{dx}\Big|_{x_c} < 0 \Rightarrow \tilde{x} < x_c$$

For $(1 - \varepsilon)^{\frac{2}{\varepsilon}} V(\bar{p}) < 1$

$$\frac{dW}{dx}\Big|_{x_c} > 0$$

and

$$\tilde{x} > x_c$$

For $(1 - \varepsilon)^{\frac{2}{\varepsilon}} V(\bar{P}) > 1$

$$\frac{dW}{dx}\Big|_{x_c} < 0$$

and

$$\tilde{x} < x_c$$

Since $\tilde{x} \geq x^*$, if the "prize"

$$(1 - \varepsilon)^{\frac{2}{\varepsilon}} V(\bar{P})$$

is sufficiently large, the competitive innovating firm will overinvest in development relative to the optimal level.

Of course, patent rights are limited in duration, and even with the patent the firm may not be able to capture the full value of the innovation, as others will invent around the patent.

These results show that for $\varepsilon < 1$, there are private incentives for the development of substitute sources of supply: however, the market-determined allocation of investment in research need not be optimal. In principle, the market distortion can be corrected. For the preceding example, this would require levying a tax (subsidy), τ, such that

$$(25) \qquad (1 - \varepsilon)^{\frac{2}{\varepsilon}} V(\bar{P} + \tau) = 1$$

or

$$(26) \qquad \tau = \left(\frac{\varepsilon}{r}\right)^{\frac{\varepsilon}{1-\varepsilon}} (1 - \varepsilon)^{\frac{3-\varepsilon}{1-\varepsilon}} - \bar{P}$$

(if $S_0 = 0$).

For $\varepsilon \to 0$,

$$\tau = 1 - \bar{P} \quad ,$$

and for $\varepsilon \to 1$,

$$\tau = \bar{P} \quad ,$$

a subsidy equal to the unit production cost of the substitute.

For efficient distribution, the market price of the substitute should remain at $\bar{P}$.

In the next section, we consider $\varepsilon > 1$ $[\alpha < 1]$, and show that the policy implications for this case can be quite different from those described above.

8.5 Competitive Development of Substitute: Monopoly-Owned Resource

This situation is entirely different from the preceding case, for two reasons. First, if the exhaustible resource is owned by a monopolist, the price would be determined by the cost of a substitute. If development of the substitute were socially managed, the substitute should be introduced if the monopolist maintained the resource price above

$$(27) \qquad \bar{h} = \bar{p}\left[1 + \frac{r}{\varepsilon}\left(\frac{\varepsilon - 1}{\bar{p}}\right)^{\frac{\varepsilon-1}{\varepsilon}} A\right]^{\frac{\varepsilon}{\varepsilon-1}}$$

for the case in which development costs are a constant, A, and $\alpha < 1$. Thus $\bar{h}$ is a limit price for the monopolist. For $\alpha > 1$, the profit-maximizing monopolist schedules production to exhaust the stock when the resource price equals $\bar{h}$. In both cases, the cost of developing and producing the substitute determines the monopolist's profit-maximizing pricing strategy. Or, more precisely, it is the monopolist's perception of the cost of developing and producing the substitute that determines the monopolist's optimal strategy.

The second difference is the incentive structure for decentralized development of the substitute. When $\alpha > 1$ and a substitute is introduced, the monopolist would lower its price to exhaust the stock when the price equals the substitute price. When $\alpha < 1$, the monopolist need only lower its price to just undercut the substitute price. This is illustrated in figures 7 and 8.

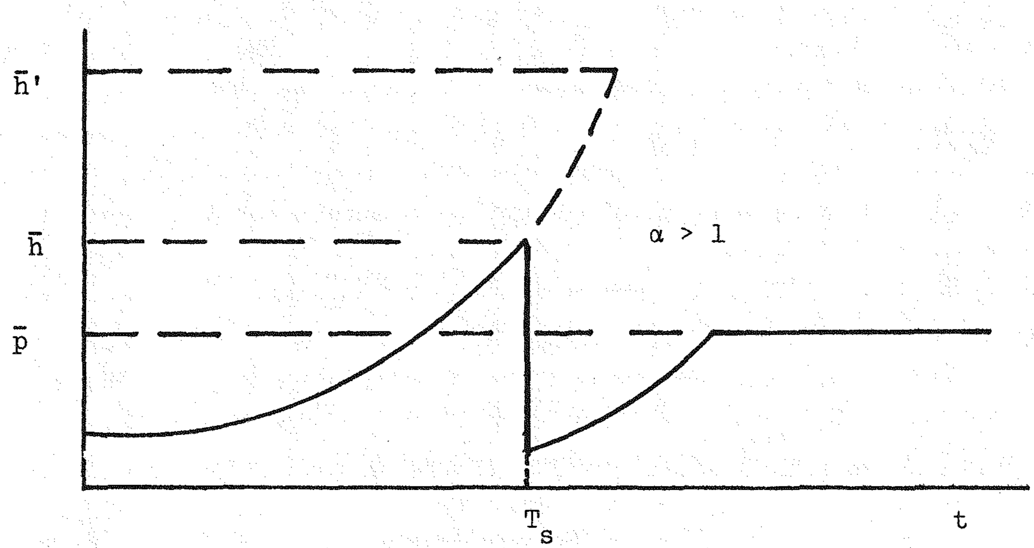


Figure 7

We assume that before the substitute is introduced, there is a backstop price of $\bar{h}' > \bar{h} > \bar{p}$. In figure 7, $(\alpha > 1)$, the substitute is introduced at T_s, and the resource price drops to clear the stock when $P_t = \bar{p}$, the substitute price. In figure 8, $(\alpha < 1)$, the price is $\bar{h}'$ until T_s, when it drops to just below $\bar{p}$.

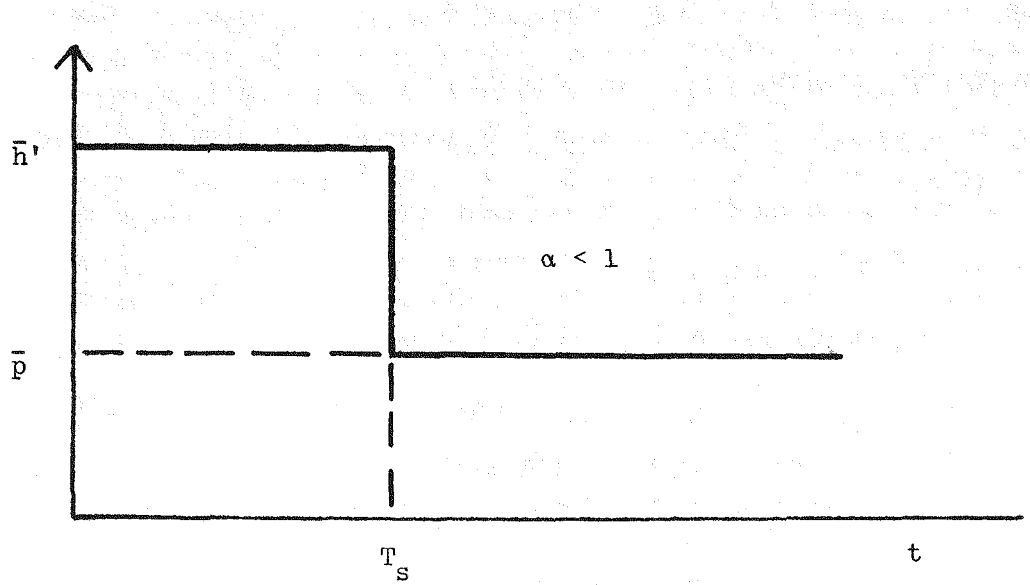


Figure 8

In both cases, introduction of the substitute must wait exhaustion of the monopolist's stock.

It may seem that it is in the interest of the consuming nation to subsidize research in order to reduce the monopolist's market power. However, if the cartel knows $\bar{p}$ and A, it can calculate the limit price $\bar{h}$. It is not in the interest of the cartel to price in excess of $\bar{h}$. Hence the consuming public should be <u>indifferent</u> with respect to development of the substitute as long as the cartel price does not exceed $\bar{h}$. In other words, development of substitutes reduces monopoly profits.

This assumes that the cost of the substitute is public knowledge. It is to the advantage of the consuming nation to convince the cartel that the cost of the substitute is low, since this lowers the cartel's expected profit-maximizing price. In an atmosphere of mutual distrust, the cartel might view information as propaganda and price above the announced cost of substitutes. Hence it may be necessary for the consuming nation to introduce the substitute (even though it may not be used) in order to police the actions of the cartel.

In general, the pricing policy will depend on the state of the substitute technology. The limit price will be high if the capacity to produce the substitute is low, or if the price of the substitute is high.

If expectations are rational, decentralized producers of the exhaustible resource (in the fringe) will invest in the expansion of capacity in such a manner that the cost of imports from the cartel and the cost of capacity is (efficiently) minimized.

The cartel will set its price aware of its incentive on the consuming nation to develop the substitute technology; and conversely the consuming nation should organize research to limit the monopoly power of the cartel, as well as to provide a source of supply alternative to the exhaustible resource. This is a complicated problem when one takes account of the reactions of each party to the actions of the other party. However, it is not difficult to determine the result if we assume that the price of the substitute depends on the actions of the consuming nation, and the cartel simply takes the current price of the substitute as given and prices accordingly. Then the consuming nation will invest in the development of the substitute in order to minimize

$$(28) \qquad \{\int_0^T P_t(h(I))q\,(P_t)e^{-rt}dt + \frac{1}{r}V(h(I))e^{-rT} + I\}$$

where

I = total investment in substitute

h(I) = trigger price of substitute

Let

$$(29) \qquad S^e = \frac{\varepsilon}{\varepsilon - 1}S_f \quad .$$

Now

$$P_t(h(I)) = P_0(h(I),S^e)e^{rt}$$

therefore the objective is

$$(30) \qquad W = \min_I \{P_0(h(I),S^e)S^e + \frac{e^{-rT}}{r}V(h(I)) + I\} \quad ,$$

with

$$P_0 = (\varepsilon - 1)\left[\frac{rS_f}{\varepsilon - 1} + \left(\frac{\varepsilon - 1}{h(I)}\right)^{\frac{1}{\varepsilon}}\right]^{-\varepsilon} \quad \text{, and}$$

$$e^{-rT} = \frac{P_0}{h(I)} \quad .$$

The optimal level of investment is inversely related to the stock remaining at time t, $S_f(t)$, as shown in figure 9

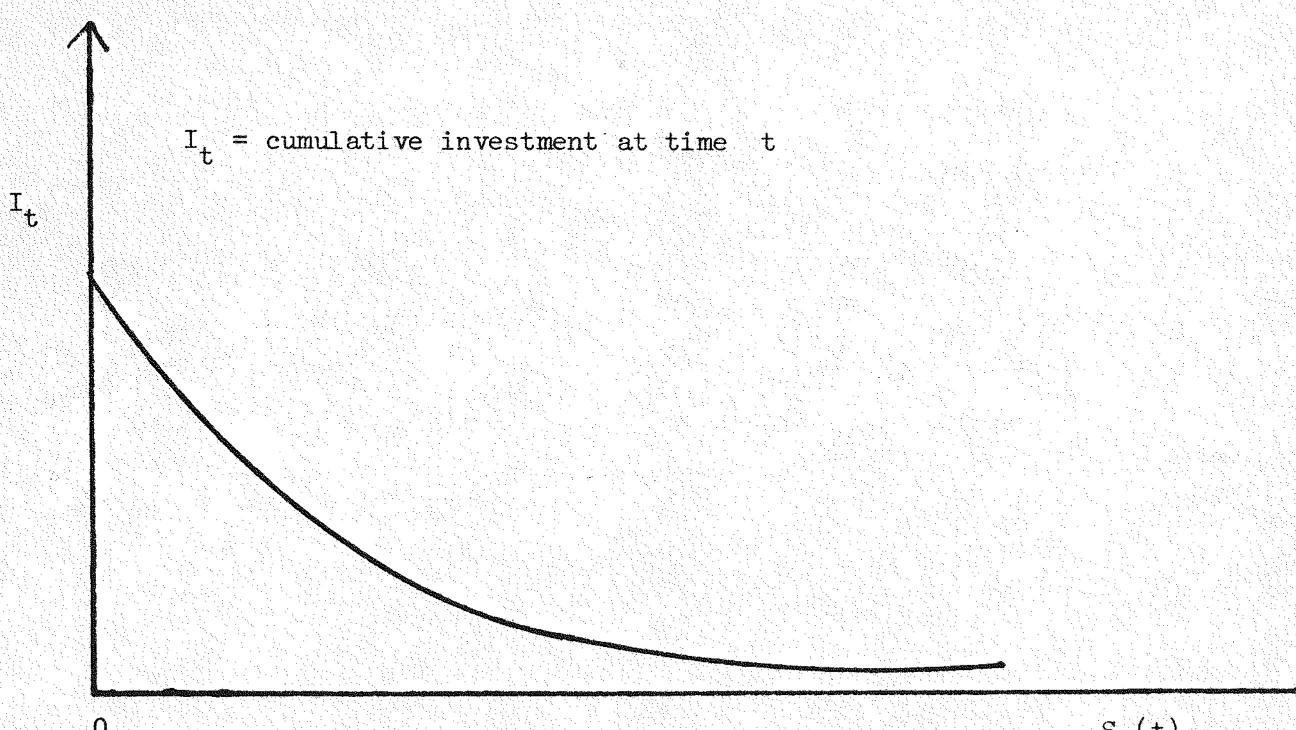


Figure 9

It is important to note that given the market power of the cartel, the value of investment in research on a substitute source of supply may be primarily its indirect effect on the cartel's profit maximizing price. In other words, the substitute may not be actually used, at least for some time. It's value is the threat of its use against the cartel.

Section 4: Footnotes

1/ The level of investment, x, represents the minimum of the present value costs of the time paths of investment that make the substitute available at time T.

2/ If cumulative resource consumption is significant during the time required to develop the substitute (see footnote 1), the efficient level of investment may be a continuous function of time.

9. Impact of Uncertainty on Resource Supply

9.1 Production Cost Uncertainty

The net social value of the substitute,

$$V(p) = \frac{\varepsilon}{r}\left(\frac{1-\varepsilon}{p}\right)^{\frac{1-\varepsilon}{\varepsilon}} \quad \text{if} \quad \varepsilon < 1$$

$$= -\frac{\varepsilon}{r}\left(\frac{p}{\varepsilon - 1}\right)^{\frac{\varepsilon-1}{\varepsilon}} \quad \text{if} \quad \varepsilon > 1$$

is convex in p. Therefore, the expected social value of a program with uncertain production cost exceeds the social value of the expected production cost. as shown in figure 1. This implies that the trigger price for an uncertain technology is <u>less</u> than the trigger price for the expected cost of the substitute.

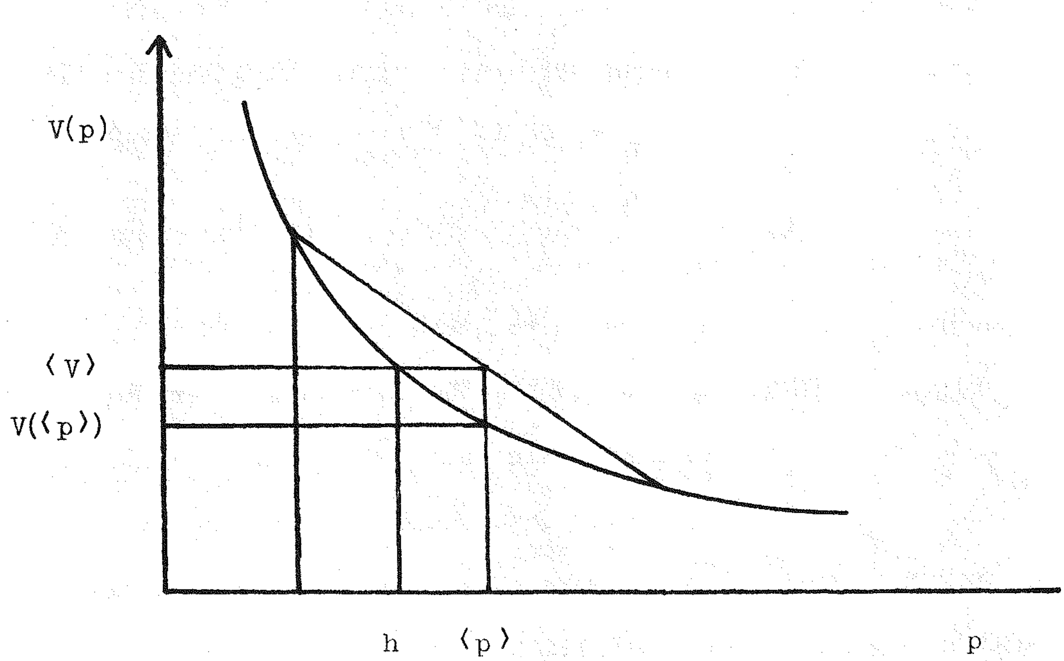


Figure 1

We define the trigger price, $\tilde{h}$, by $V(\tilde{h}) = \langle V \rangle$. That is, $\tilde{h}$ is the cost that yields the expected value of the development program. Since

$$\langle V \rangle > V(\langle p \rangle) \quad ,$$

it follows that

$$\tilde{h} < \langle p \rangle \quad .$$

The consuming nation could be better off if the cost of a substitute were uncertain than if the substitute were known to have a cost equal to the mean of the cost distribution. For these results to hold, it is necessary that both the cartel and the consuming nation share the same information about the substitute technology.

If it were possible to prevent the cartel from gaining information about the substitute technology, the consuming nation could benefit from a research program. Information about the cost of developing the substitute would be valuable to the consuming nation if it could be kept secret. For example, if research demonstrated that shale oil would be very expensive, then research efforts could be allocated to, say, coal liquefaction. If this information were communicated to the cartel, the cartel would revise its calculation of the limit price upwards. If the information were kept secret, the cartel would not revise its limit price and the consuming nation would be better off. If the research showed that shale oil would be cheap, the consuming government would elect to make this information public and the cartel would be induced to lower its limit price.

9.2 Randomized Prices and Incentives For Instability

Copper is an example of an exhaustible resource produced by a small number of large firms for which there is an active futures market, and, given such an institutional structure one might have expected a stable price trend for copper. The firms are large enough to affect the price, copper is conveniently storable, and if the producers wanted to they could doubtless speculate in the forward market to stabilize the price. And yet the price is very instable. We saw in section 9.1 that consumers gained if prices were randomized around an unchanged mean, and it is natural to ask whether a monopolist (or cartel) could exploit this consumer benefit of price instability to raise his own profits.

Let us examine the case where the monopolist faces a trigger price $\bar{h}$, below his unconstrained profit maximizing price. If he sets a stable price higher than $\bar{h}$ the consuming country will make an irreversible investment decision to produce a substitute, as discussed in section 4. If the consumers can correctly predict the monopolist's behavior and if the monopolist is well informed about the trigger price then the monopolist's best non-random strategy is to set the price marginally below the trigger price and thus deter investment in producting the substitute. We saw that the consumers have an incentive to increase the dispersion of possible operating costs of producing the substitute about any given point estimate, since this lowers the trigger price, just as they have an incentive to lower the monopolist's perception of the estimated cost of the substitute. The monopolist has a similar incentive to create uncertainty, for similar reasons.

Let us define $\bar{h}$ to be the trigger price calculated on the assumption that futures prices are known with certainty, so that the consumer will invest if he predicts the monopoly price to be constant and higher than $\bar{h}$. If the consumer is led to believe that the future monopoly price is random with an expectation higher than $\bar{h}$, then he may prefer to buy from the monopolist at the varying price rather than invest in the substitute, for since his indirect utility function $V(p)$ (which measures the benefit of buying at price p) is convex in prices, if p is random

$$EV(p) > V(Ep) \quad .$$

If $\hat{p}$ is such that $V(\hat{p}) = EV(p)$, $\hat{p} < Ep$, then the monopolist can deter investment if he can arrange $\hat{p} < \bar{h}$.

The monopolist will favor any costless method of randomizing price expectations about an unchanging mean, and will also have an incentive to lower the forecast price. Consumers will be aware of the latter incentive, and may be able to correct for this bias in their forecasts, but it is not obvious what they can do to reduce the randomness in the forecast price. If it is sufficiently cheap, we would expect the monopolist to be a destabilizing speculator in forward markets, or even to surpress forward markets and disseminate misleading information. A cartel will spread rumors of its imminent collapse and may maximize the publicity given to internal difficulties (providing these are not misinterpreted by the cartel members). The arguments above suggest that these rumors should not necessarily be taken at face value.

Speculating in futures markets may be a relatively cheap method open to the monopolist of destabilizing price expectations, for the volume of trade conducted at these random prices need not be large. The monopolist can also destabilize the spot price, and, in the absence of a futures market this is the most obvious way he can persuade consumers that prices are random. Would it pay the monopolist to randomize the spot price?

The question of the costs and benefits of price stabilization versus price instability has a long history, dating back to World War II when Waugh[1/] argued against schemes to stabilize prices, and including Samuelson's [2/] counter claim that the consumer does benefit from feasible price stabilization. We still lack a complete analysis of price stabilization (see Newbery[3/] for recent extensions of the analysis) and it is worth pointing out that all the results so far apply to competitive markets, but it is easy to see why in a technologically certain world with perfect competition price stability, which is, an hypothesis feasible, is also desirable. Figure 2 graphs the sum of consumer surplus and profits, $W(p)$, as a function of price for the case where demand is isoelastic and marginal production costs are constant.[4/]

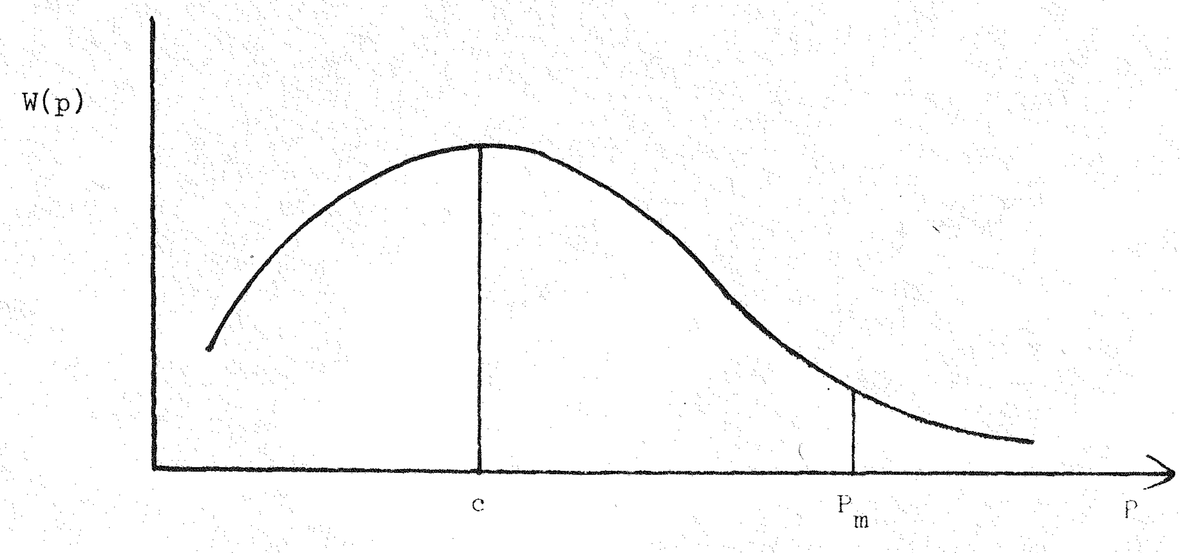


Figure 2

If demand is $q = ap^{-\alpha}$ where α is elasticity of demand, then consumer surplus

$$S(p) = \int_{p}^{\bar{p}} q\,dp = \frac{a}{1-\alpha}[\,p^{1-\alpha}]_{p}^{\bar{p}}$$

where $\bar{p}$ is a reference price.

Profit is $R(p) = qp - qc - k = ap^{-\alpha}(p - c) - k$ where c is the constant marginal cost, k is fixed cost. Net social surplus is

$$W(p) = S(p) + R(p) = a\{\frac{\alpha}{\alpha - 1}p^{1-\alpha} - cp^{-\alpha}\} - K$$

with K another constant.

$W(p)$ has its maximum at the competitive price $p = c$ and is strictly concave there. This is of course a perfectly general result, independent of the particular demand and cost functions. Artificially induced price instability will lower the valueof $EW(p)$ below $W(Ep)$ at the competitive equilibrium, hence feasible stability is desirable.

The reason for thinking that this argument may not hold for a monopolistic equilibrium is that $V(p)$ is convex at p_m, the monopoly price.

$$p_m = \frac{\alpha}{\alpha - 1} \cdot c \quad \text{if} \quad \alpha > 1$$

$$\frac{d^2W}{dp^2} = a\alpha^2 p^{-(\alpha+2)}\{p - \frac{1+\alpha}{\alpha}c\}$$

$$\frac{d^2W(p_m)}{dp^2} = ap_m^{-(\alpha\pm1)} > 0 \quad .$$

Therefore feasible price randomization around an unchanged average price near the monopoly price raises the sum of producer and consumer surplus, suggesting that feasible destabilization might pay. This indeed is the case in some examples, but it does not seem to be generally true that price destabilization pays. The reason is that for a random price strategy to yield the same consumer benefit as a stable price strategy, the average quantity supplied must be greater, for consumer benefit is a concave function of quantities consumed. This raises the average cost of supply to the monopolist. In other words, it may be misleadiqg to focus on prices instead of quantities. In the constant elasticity case V as a function of q rather than p has the form

$$W(q) = \frac{a^{\frac{1}{\alpha}}\alpha}{\alpha - 1} q^{1-\frac{1}{\alpha}} - cq$$

and

$$\frac{d^2W}{dq^2} = -\frac{a^{\frac{1}{\alpha}}}{\alpha} q^{-(1+\frac{1}{\alpha})} < 0$$

so W is concave in q over the entire range of values of q. Consider the constant elasticity case with zero prodiction costs for the monopolist. His problem is to choose p to maximize $Ep^{1-\alpha}$ subject to $ES(p) \geq \bar{S}$, or $Ep^{1-\alpha} \leq \bar{h}^{1-\alpha}$ where $\bar{h}$ is the trigger price. The monopolist cannot earn more than $\bar{h}^{1-\alpha}$, and if his production costs are nonzero he maximizes profits by minimizing average supply to yield this revenue, that is, by stabilizing the price.

However, under at least some conditions it is possible for the monopolist to increase his profits by randomizing prices. In figure profits are not concave in prices at the limit price and the convex combination of prices p_1 and p_2 with an average of $\bar{h}$ will yield higher profits than any single price $p < \bar{h}$.

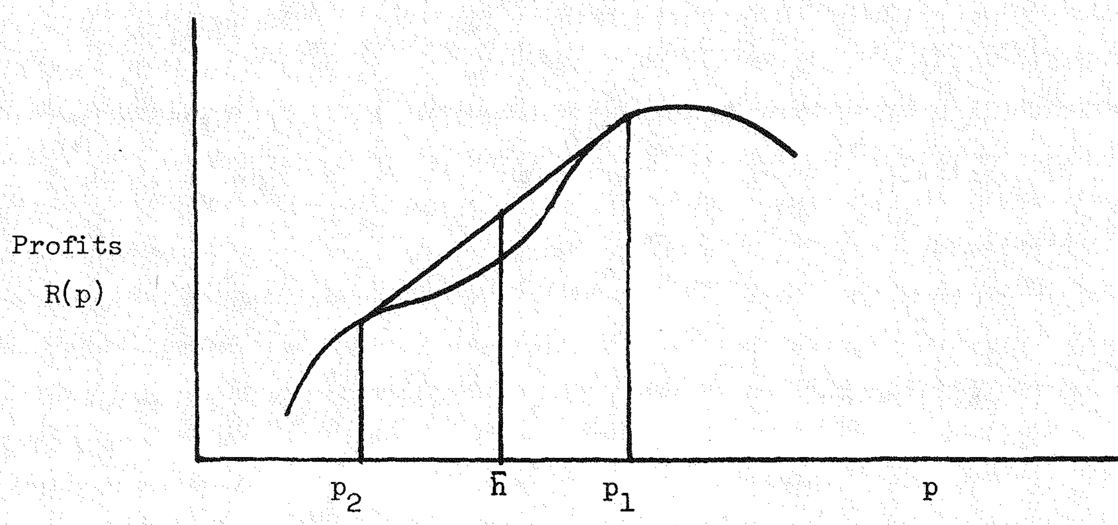


Figure 3

Actually the monopolist can do even better than this for two reasons. By randomizing price the monopolist can charge a higher average price than $\bar{h}$, and by simultaneously choosing p_i and Π_i, the probability of charging p_i, the monopolist can do better than by choosing first p_1, p_2 and then Π_1, Π_2. This will be demonstrated in the example below, after setting out the problem formally.

Let $V(p)$ be indirect utility, $R(p)$ profits, then the monopolist chooses prices p_1, p_2 $(p_1 > \bar{h} > p_2)$ with probability $\Pi_1 = \Pi$, Π_2 $(= 1 - \Pi)$

to maximize expected profits $R = \Sigma\Pi_i R(p_i)$ subject to the constraint $\Sigma\Pi_i V(p_i) \geqq V(\bar{h}) = \bar{V}$ say.

Then

$$\Pi_i = \frac{V_j - \bar{V}}{V_2 - V_1} \quad , \quad j \neq i$$

$$\frac{\partial\Pi}{\partial p_i} = \frac{V_i'\Pi_i}{V_2 - V_1}$$

$$\frac{\partial R}{\partial p_1} = \frac{\partial\Pi}{\partial p_1}(R_1 - R_2) + \Pi R_1'$$

and

$$\frac{\partial R}{\partial p_i} = \Pi_i\{kV_i' + R_i'\} \quad , \quad i = 1, 2$$

where

$$k = \frac{R_1 - R_2}{V_2 - V_1} > 0 \quad , \quad R_i' = \frac{\partial R(p_i)}{\partial p}$$

If the marginal utility of income λ is independent of p, then $V_i' = -\lambda q_i$, and λ can be taken as unity.[5/] If also there is an interior maximum

$$\frac{1}{q_1}\frac{\partial R_1}{\partial p} = \frac{1}{q_2}\frac{\partial R_2}{\partial p}$$

In particular, if $R = pq - cq$, and $\alpha = -(p/q)(\partial q/\partial p)$ is the elasticity of demand, then

$$\alpha_1(1 - \frac{c}{p_1}) = \alpha_2(1 - \frac{c}{p_2}) \quad .$$

Since $p_1 > p_2$, $\alpha_1 < \alpha_2$, so the demand curve must become more elastic at lower prices for randomization to pay, a condition which is empirically plausible. In the special case of zero opportunity cost, $c = 0$, $\alpha_1 = \alpha_2$ is a necessary condition for a maximum.

The simplest example which illustrates the gains from randomization is shown in figure 4.

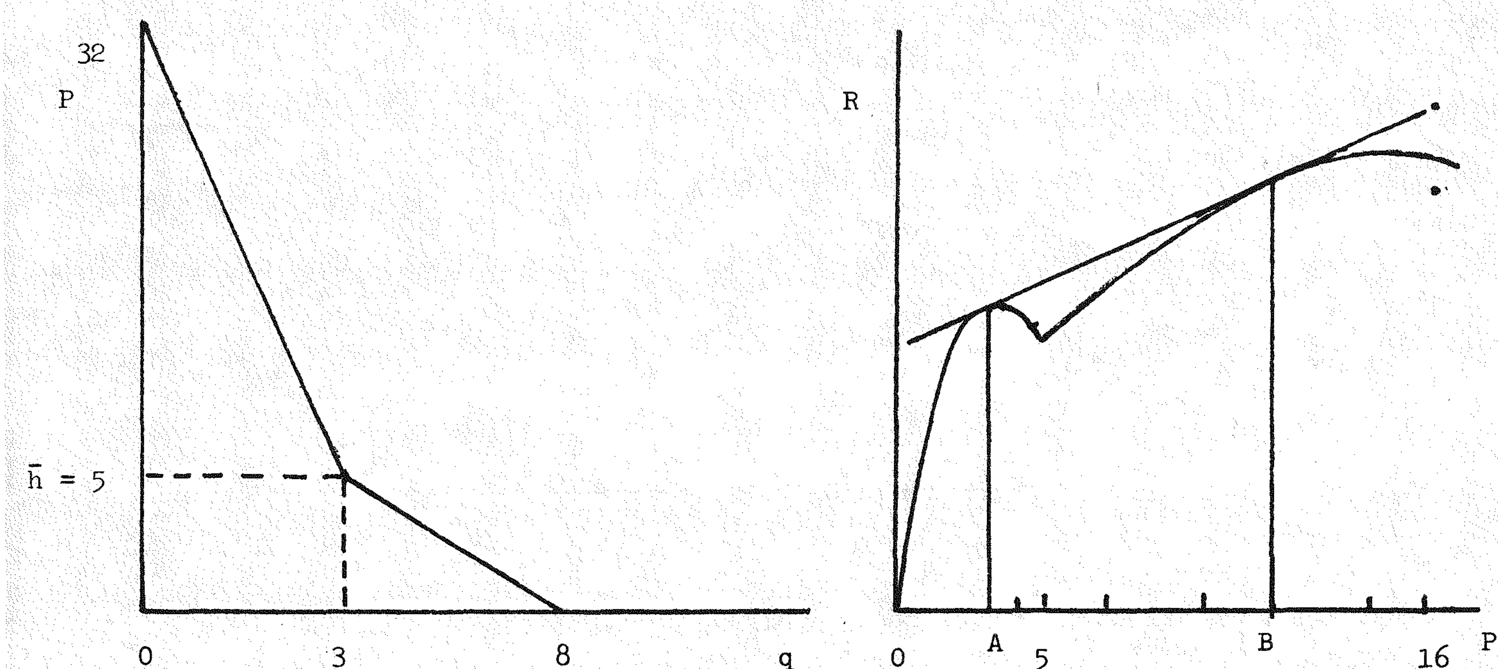


Figure 4

The demand curve facing the monopolist is

$$p = 32 - 9q \qquad q \leq 3$$
$$= 8 - q \qquad q \geq 3$$

and the trigger price $\bar{h} = 5$. Costs are zero, and the profit function is bimodal, having maxima at $\bar{p} = 4$, $p = 16$. In the absence of the trigger price the monopolist would set price $p = 16$ and earn profits 28 4/9, whilst if he is constrained to it a constant price no higher than $\bar{h} = 5$, then he sets $p = 4$ earning 16. A convex combination of these two prices with $\Pi_1 = 1/9$ yields profits 17 31/81 and an average price of 5 1/3 and the consumers still prefer to buy from the monopolist. However, if the monopolist chooses prices OA, OB to convexify the profit function (see figure 3) then $p_1 = 15$, $p_2 = 1\ 2/3$, and with $\Pi_1 = 7/18$, average prices rise to 5 17/27 and profits increase to 17 38/81. But if the monopolist solves the equations $\partial R/\partial p_i = 0$, where $p_i = A_i - a_i q$, then $p_1/p_2 = A_1/A_2 = 32/8 = 4$, then the optimum requires $p_2 \cong 2.9$. If $p_2 = 3$, $p_1 = 12$, $\Pi_1 = 8/21$, the average price rise to 6 3/7, and profits to 19 4/9. Thus the best randomized strategy is considerably better than naive randomizing strategies (of choosing combinations of the two local profit maxima, or convexifying the profit function), and is itself considerably better than a constant price strategy which yields some 80% of the average random profits.

This result appears to be quite new and very unexpected. We have shown that under some circumstances it will pay a monopolist to randomize his price and destabilize the market, and that more generally he benefits by increasing the uncertainty about future prices where consumers face irreversible investment decisions. The result should be distinguished from two other cases in which random prices may be profitable. The first is where both consumer and monopolist are uncertain about the other's best strategy, and each chooses a mixed strategy as the best game theoretic solution. Here the monopolist knows exactly what the consumer's response will be, and the consumer knows the monopolist's strategy. In the language of oligopoly we have a determinate Stackleberg equilibrium rather than indeterminacy. The second is Salop's noisy monopolist, who randomizes price discrimination.[6/] Here there is an intertemporal price discrimination, but the consumers are assumed homogenous, have no search costs, and are not separated into distinct submarkets.

Finally, we can ask what the government's policy towards a randomizing monopolist should be. If storage costs are high so buffer stocks are precluded, then the government can only stabilize prices by regulation, with the regulatory agency making profits or losses as prices fluctuate below or above the average. The strategy would lower the monopolist's average profit, but will only ruin consumer welfare if, as in the numerical example, the profit function were bimodal. Since this is not necessary for randomization to be profitable, there is no guarantee that consumers will benefit from regulation. Moreover, if the price instability were partly or wholly caused by environmental uncertainty, then consumers stand to lose by having the price regulated.

If, on the other hand, storage costs are low consumers <u>may</u> benefit from creating a buffer stock (though the monopolist's behavior would then change, and we would expect him not to find randomization profitable). The distribution of benefits from operating a buffer stock in the presence of environmentally caused supply uncertainty is discussed at some length in

Newbery.[7/] We conclude that whether price stabilization is desirable or not will depend on the source of the instability, whether buffer stocks can be carried sufficiently cheaply and such factors as the elasticity of decision and the nature of the risk, whose quantitative importances are analyzed in Newbery.[8/]

Section 9: Footnotes

1/ Waugh, F. V. [1944], "Does the Consumer Benefit from Price Instability?" <u>Quarterly Journal of Economics</u> 58, pp. 602-14.

2/ Samuelson, P. A. [1972], "The Consumer Does Benefit from Feasible Price Stability," <u>Quarterly Journal of Economics 86</u>, pp. 476-93.

3/ Newbery, D. [1976a], "Price Stabilization with Risky Production," IMSSS Working Paper, Stanford University.

4/ See Newbery [1976b] for a general analysis.

5/ For this to be true when income is y, $V = V(p,y)$, $\lambda = V_y$

$$V_p = -qV_y, \quad V_{py} = \frac{-qV_y}{y}\left(\frac{y}{q}\frac{dq}{dy} - \left(\frac{-V_{yy}\,y}{V_y}\right)\right) = 0$$

i.e. the income elasticity of q is equal to the relative risk aversion.

6/ Salop, S. [1973], "The Noisy Monopolist: Imperfect Information, Price Dispersion and Price Discrimination," Federal Reserve Board.

7/ Newbery, D. [1976a], "Price Stabilization with Risky Production," IMSSS Working Paper, Stanford University.

8/ Newbery, D. [1976b], "Feasible Price Stability May Not Be Desirable," IMSSS Working Paper, Stanford University.

9/ Stiglitz, J.S. [1972], "Taxation, Risk Taking and the Allocation of Investment in a Competitive Economy," in M. C. Jensen (ed.) <u>Studies in the Theory of Capital Markets</u>, Praeger, New York.

10/ Of course there are other sources of uncertainty in the development of substitutes, in the bahavior of the cartel, and more generally about the level of future income and hence demand. The effect of uncertainty about the cartel's behavior is discussed later, whilst some of the other issues have been addressed in earlier sections.

11/ Parenthetically we should note that the increased cost of providing the flexibility may take the form of a higher average rate of consumption of energy, and hence a greater average dependence on and vulnerability to foreign suppliers. But it may also take the form of more flexible capital equipment and not increase average dependence. We hope to look at both cases to see if they systematically differ.

10. The BTU Tax

The incidence of a BTU tax would depend on the manner in which the tax is administered. There is no single homogeneous source of energy. Energy sources differ in ease of transportation and handling, combustion temperature, emissions, alternative uses for chemicals, etc. Consumers differ in their preferences toward these attributes of energy sources. Thus the incidence of a tax on the best output of an energy source would be the aggregate of responses by diverse consumers to the diverse sources and uses of energy resources.

This complexity immediately raises concern over the efficiency of such a tax. If the market price of an energy resource differs from its social value, it is unlikely that this difference would be apportioned equally over the many different sources of energy and their uses. Furthermore, consumers may value differently the extent to which the market price differs from the socially optimal price. We shall discuss some of the general effects of a BTU tax and raise some of the issues associated with its lack of specificity in what follows.

10.1 Intertemporal Shifting of the BTU Tax

In a static analysis, the incidence of a tax depends on own and cross-elasticities of demand and supply. However, the central issue in the allocation of an exhaustible resource is the trade-off between consumption now versus consumption in the future. Consequently, analysis of the incidence of a tax on an exhaustible energy resource should address the effect of the tax on the pattern of energy use over time.

We neglect for now the difficulties of determining the tax incidence on different grades and forms of energy, and assume there is only one energy good, e.g., oil of a particular viscosity, sulfer content, etc. This may be extracted from an exhaustible stock at zero cost, or produced without limit at cost $\bar{P}$.[1/] We know from the discussion in the first interim report that in the absence of uncertainty the price trajectory corresponding to an efficient allocation of the resource would be as shown in figure 1. Furthermore, the allocation generated by a perfectly competitive economy is efficient.

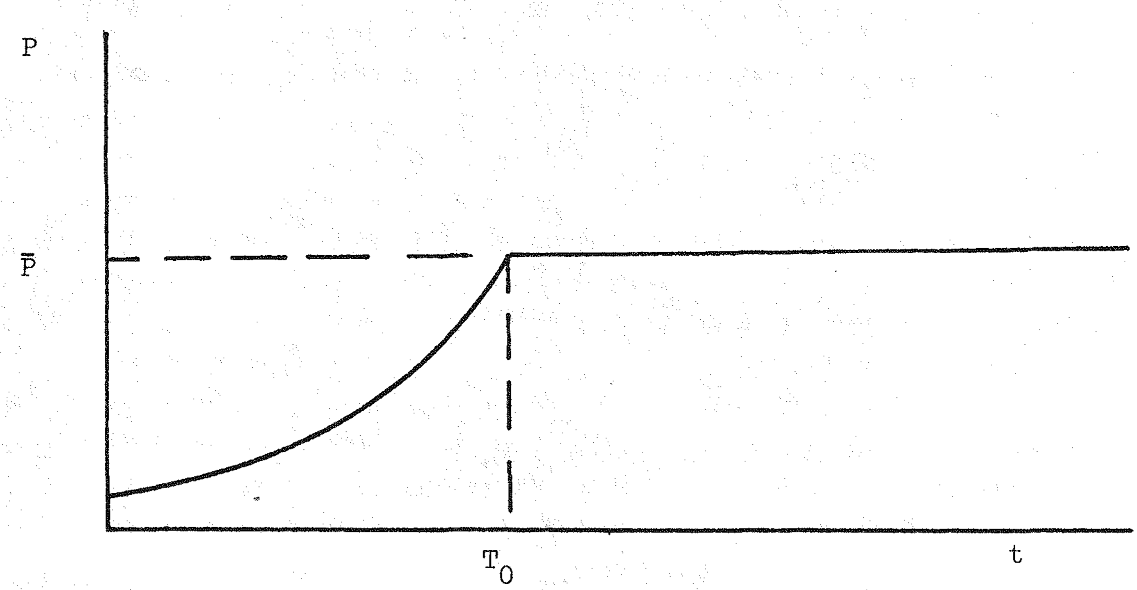


Figure 1

A constant absolute BTU tax will, among other things, distort the intertemporal pattern of energy use in a closed economy. An equilibrium will still require that the net rent on exhaustible resource stock rise at the market rate of interest. However, with a BTU tax, the price consumers pay will differ from the price paid to the producers of the exhaustible resource. Since there is only one energy good in this example, a BTU tax is equivalent to a tax per unit of the resource (e.g. per barrel of oil).

Let x denote the tax per unit of energy, $V(t)$ the rent on a unit of the exhaustible resource, and $P(Q(t))$ the price consumers are willing to pay for Q units of energy. The allocation will be independent of whether producers or consumers pay the tax, since in either case

$$P(Q(t)) = V(t) + x \quad .$$

Suppose only the exhaustible resource is taxed. The equilibrium price trajectory after imposition of the tax is determined by

$$\frac{\dot{V}(t)}{V(t)} = r \quad , \quad \text{and}$$

$$\int_0^T Q(t)\,dt = S_0 \quad ,$$

where S_0 is the size of the resource stock,

$$Q(t) = P^{-1}[V(t) + x] \quad ,$$

and T is defined by

$$P(T) = \bar{\bar{P}} \quad .$$

The after tax price trajectory may look as shown in figure 2.

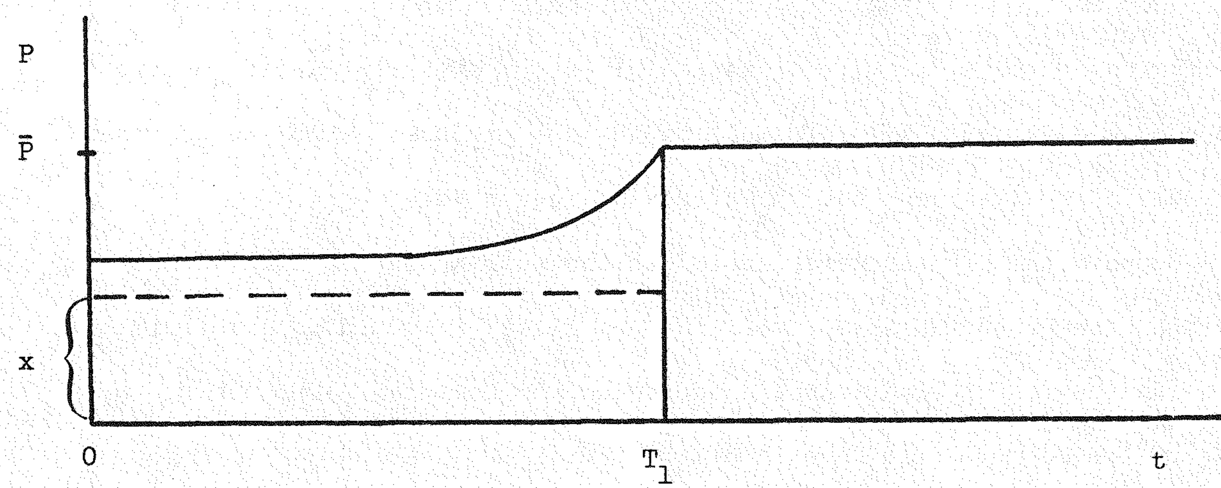


Figure 2

The tax induces an allocation that is too conservative relative to the efficient allocation. In particular, the time at which the resource is exhausted, T_1, exceeds T_0 in figure 1.

If consumers pay the tax, the price producers receive is $P(t) - x$. If producers pay the tax, they receive $P(t)$ per unit of energy. The tax increases the cost to consumers of energy from the exhaustible resource, and lowers rents on the resource. Thus the tax may have important distributional effects. Indeed, an argument for an energy tax is to reduce windfall profits that accrue to owners of an exhaustible resource. However, the BTU tax is not a profits tax. A pure profits tax would tax only the rent component of the supply price, and hence leave supply and demand unchanged. The BTU tax essentially increases the cost of producing the exhaustible resource. This changes the private rate of return on the resource and alters the market allocation.

In this example, the substitute is not taxed. Thus there is no change in this market, except that the substitute is introduced at a later time.

When energy is purchased for consumption, it may not be feasible to identify its source -- i.e., the exhaustible stock or the produced substitute. In such a case it would be preferable to tax the producer, since to do otherwise would effectively tax the substitute source as well.

Suppose the BTU tax were carried to the extreme and applied to both sources of energy at the same rate. In that event, it would be immaterial whether the tax is applied to the producer or the consumer. The after tax price trajectory would be as shown in figure 3.

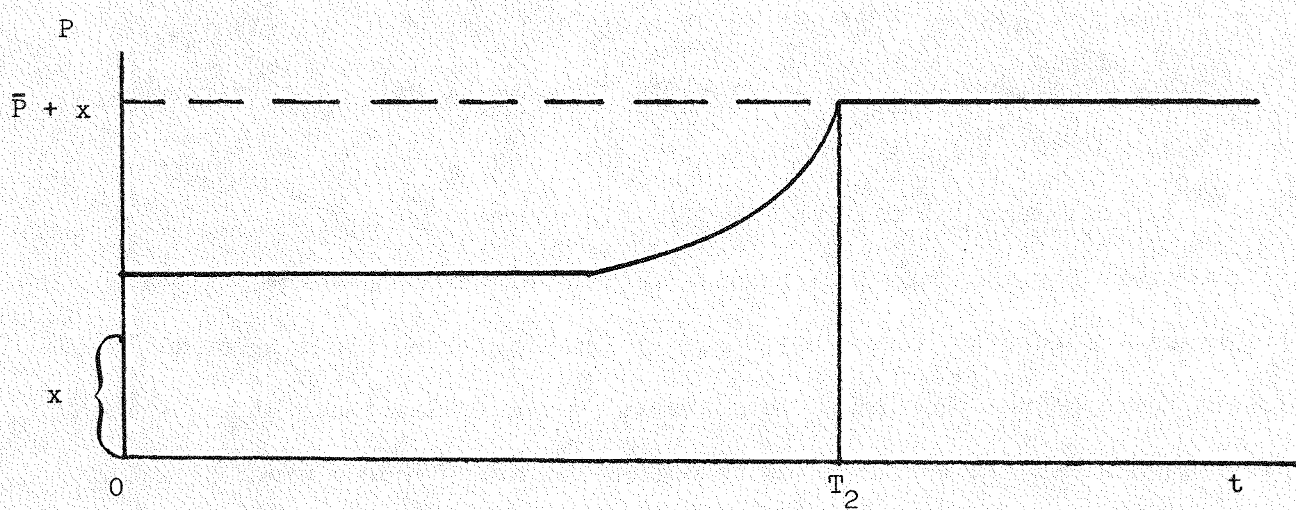


Figure 3

Taxation of the substitute source increases the private rent on the exhaustible resource, relative to the previous example. Consumer costs are increased, and the substitute is introduced at a date later than T_1.

There is a special case worthy of consideration. Suppose there are fixed costs in the development of substitutes, and the substitute is produced under the protection of a patent, so that $\bar{P}$ exceeds the production cost. A tax on the substitute would decrease revenues. If, as a result, total discounted revenues fell short of fixed costs, the substitute might not be introduced at all. This should dramatize the importance of tax specificity. A general BTU tax could have perverse effects. By reducing the incentives to develop substitute sources of supply, the BTU tax can exacerbate energy scarcity.

It should be clear that in a perfectly competitive economy with no uncertainty, the BTU tax can have at best a positive impact on the redistribution of income. However, this would be at the cost of a less efficient allocation of energy over time. Of course the world is not predictable and the economy is not perfect. We consider below the effect of a BTU tax when there are market distortions. But first let us examine a difficulty that arises when we relax the assumption of a single energy good.

10.2 The Choice of Tax Base

When there is only one energy good, the BTU tax can be considered as either a tax on BTU's or a tax on units of the good. When we consider different energy forms, and different grades of energy sources, the choice of what to tax is more difficult. Suppose the tax is on the heat energy actually produced by the energy source. This would be quite difficult to monitor, but let us neglect that problem. In this case, more efficient use of a resource would imply a higher tax per unit of the energy source (e.g. per barrel of oil). Thus the tax would discourage efficient energy use.

Suppose, instead, that the tax is added to the price of each source of energy independent of its use. Clearly, the tax would have to differ for each type and grade of energy resource, reflecting the characteristic heat value of the source. Presumably the tax would be lower on unprocessed ores and poorer grades of a resource. Thus there is a potentially serious accounting problem. If the tax rate on different energy forms deviated from a uniform rate on a BTU basis, there will be an incentive to substitute fuels that are taxed at the lower rate. Furthermore, the tax rates would have to be revised to reflect technical progress in the use of specific fuels. The multiple uses of fuels present additional difficulties. Should oil used to manufacture drugs or synthetic proteins be taxed at the same rate as oil used for space heating?

10.3 Imperfect Markets

Assuming the problem of defining the tax base has been brought under control, let us consider the incidence of a BTU tax in an economy with sources of market failure. We will argue that market imperfections often justify corrective taxes and subsidies. Unfortunately, the BTU tax is not specific enough to improve the market allocation. The opposite extreme of a tax or subsidy for every source of market distortion would be prohibitively expensive to administer. The art of regulation is to choose the set of instruments that achieves the greatest welfare gain net of administration costs.

In the following pages we describe the impact of the BTU tax on the sources of intertemporal allocative distortions identified in the first interim report.[2/]

10.3.1 Imperfect Competition

We have emphasized in this report that the magnitude of the distortion from imperfect competition depends on the size and production capacity of actual and potential competitors. However, the direction of the distortion caused by imperfect competition is unambiguous. Price-setting behavior results in a allocation that is too conservative. This implies that in imperfectly competitive markets the current price is higher than the efficient price. Hence a BTU tax only magnifies the monopoly distortion.

The tax may be effective in reducing monopoly or cartel revenues. However, as we have shown in the preceding section, the tax must be selectively applied to be effective in reducing monopoly rents. If applied to all sources of energy, monopoly rents may not be reduced significantly, while consumers would be worse off. If applied only to the resource supplied noncompetitively, monopoly rents would be decreased.

Power of taxation is clearly limited in open economies. A consuming nation can tax its consumers, but its ability to tax foreign producers or consumers is limited. The tax would reduce consumption in the country of origin, and lower the price for other consumers. To effectively tax monopoly revenue, consumer countries would have to coordinate tax policies. Such a union would be quite fragile, since each country would be tempted to break away and enjoy the low market price resulting from the coordinated tax policies of the other countries.[3/]

10.3.2 Uncertainty

There are a number of issues to consider under the general heading of uncertainty. We will begin with what appears to be a persuasive, though not convincing, argument for the BTU tax.

Forecasting Error

We have shown that the efficient price of an exhaustible resource contains a component that is the factor cost of extraction, and a component that is the scarcity value of the stock. The later rises over time at the market rate of interest. Thus, on the average the market price increases at an increasing rate over time (see figure 4).

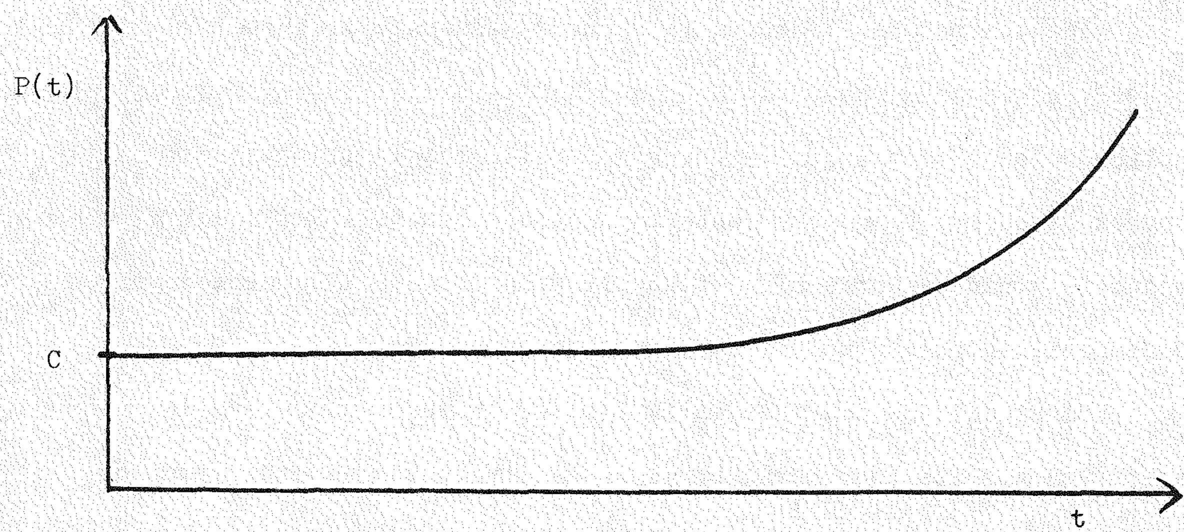


Figure 4

If consumers simply extrapolated future rates of price increase on the basis of past rates, their forecasts of future prices would be systematically low. This may be compounded by stochastic price variations. Since consumers expect future prices that are too low, current consumption will be excessive. For example, consumers may invest in energy intensive machinery that is not profitable given the true price trajectory. Yet the variable costs of the machinery may be low enough to continue operation at high fuel prices.

A tax could conceivably correct this distortion. However, there are important caveats. First, the argument about price extrapolation is not very convincing. One could make a case that consumers over-react and expect future prices that are too high. Second, the tax must be linked to the scarcity value of the energy source. This means it must be different for alternative energy sources. Furthermore, a constant tax will not correct the forecasting error. The tax must vary over time, first increasing and then decreasing as the rate of energy price increase stabilizes.

Imperfect Capital Markets

The absence of insurance may imply that the market rate of discount is high relative to the socially optimal rate (although this is a debatable point). The high rate of discount implies that exhaustible stocks are exploited at too rapid rates. The BTU tax at a constant level causes the market allocation to be more conservative.

Once again, there are problems with such a tax. The tax would have to be selectively applied on the basis of resource scarcity. Hence a general BTU tax would be inefficient. More important, if the market interest rate is too high, it is too high for all economic allocations. Therefore, a much better alternative is the proper exercise of monetary policy to reduce the market rate of interest to its optimal level.

Security

The market can adapt to supply interruptions provided market prices are permitted to reach their equilibrating values. In the event of an embargo or war, price rationing is likely. For this reason, the market may maintain insufficient reserves to meet such contingencies and the rate or resource utilization may be excessive.[4] (There is also the possibility of market bias resulting from incomplete markets and differentiated producers and consumers. The direction and magnitude of this bias has not been determined as yet.)

A tax may be appropriate to correct the bias resulting from supply insecurity and rationing. The tax must be related to the degree of riskiness in supply. In fact, what is needed is not a tax on supply from an uncertain source, but rather a portfolio tax on imports. For example, suppose oil is imported in equal amounts from countries A and B, both of which are risky sources. Yet suppose A and B are mortal enemies, if A interrupts supply, B will not, and conversely. Clearly, this negative correlation reduces total risk.

10.4 Conclusion

This has not been an exhaustive analysis of the BTU tax. There are many other factors to consider, particularly with regard to administration of the tax. However, we hope we have demonstrated the potential pitfalls of such a tax, and the desirability of more carefully designed measures.

Section 10: Footnotes

1/ The assumptions of zero extraction cost and finite cost of producting a perfect substitute can be relaxed with little change in the analysis of the tax incidence.

2/ This neglects some important sources of static distortions such as pollution externalities. Similar comments apply to the effect of the BTU tax on the static allocation, although this is not the focus of this study. The BTU tax does not account for the different environmental impacts of alternative energy sources, and thus fails to improve substantially upon the market allocation of pollution.

3/ On the other hand, a tax on energy imports would be a very simple tax to administer. Since many countries already tax energy imports, a unified tax policy for consuming nations would not require a change in policy for all consumer governments. Furthermore, once a tax is implemented, it becomes a source of revenues that governments might find difficult to relinquish.

4/ Consumers have an incentive to maintain reserves. However, there are economies of scale in storage that would not be fully exploited by decentralized consumers.