THESIS

THE WELFARE EFFECTS OF A MARKET ALLOCATION OF AN EXHAUSTIBLE RESOURCE

Submitted by

Frank A. Ward

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall, 1977

COLORADO STATE UNIVERSITY

Fall 1977

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY FRANK ALAN WARD

ENTITLED THE WELFARE EFFECTS OF A MARKET ALLOCATION OF AN EXHAUSTIBLE RESOURCE

BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Committee on Graduate Work

Robert E. Gaines

Dwight M. Blood

[illegible signature]

[illegible signature]

Adviser

ABSTRACT OF THESIS

THE WELFARE EFFECTS OF A MARKET ALLOCATION OF AN EXHAUSTIBLE RESOURCE

In recent years, cost-benefit theorists have developed "net benefit" measures of welfare change attributable to shifts in the allocation of flow resources. Presumably, such welfare-change measures have been developed as an attempt to minimize the wastage of resources on unsound projects. However, to the author's knowledge, no such welfare-change measure has been developed to rank alternative allocations of an exhaustible resource. This dissertation attempts to devise such a measure. The measure is developed in three steps.

The first step (Chapter two) is an explanation of how a free market allocates the exhaustible resource over time. Inquiry is made as to how the time path of extraction is affected by changes in (1) production costs, (2) total known supply, (3) the costs of a substitute technology, and (4) the discount rate. Knowledge of the allocation explored in this step is important, because, once determined, the market outcome can be compared to some appropriately-defined efficiency norm.

The second step (Chapter three) develops an efficiency norm as a basis for determining whether the market depletes the exhaustible resource too quickly, too slowly, or at the right rate.

The third step (Chapter four) of this dissertation is an attempt to develop a measure which quantifies how well the market's time-use of the mineral approaches an efficient allocation. The methodology used is the development of a measure of welfare change. Specifically, this measure is designed to ascertain the net change in

benefits attributable to changes in either of two generalized distortions relevant to the market for an exhaustible resource.

At all three steps, this dissertation draws from and extends the theory of exhaustible resources. The first step is an extension and refinement of the comparative statics of competitive mining theory. At the second step, the optimality properties of a market allocation over time are examined. At the third step, the degree to which the market breaks down is the subject of concern. Specifically, an extension of currently-accepted welfare-loss theory is developed and made applicable to the exhaustible resources sector.

The results of this dissertation are that, indeed, such a welfare-loss measure can be quantified. By incorporating (1) the effect of a change in a market distortion on the private profit-maximizing output path of each of $\underline{n}$ mining firms, and (2) the effect of these $\underline{n}$ output path changes on aggregate total discounted net benefits, a welfare-loss measure is developed. The measure can, in principle, rank alternative allocations of an exhaustible resource on the basis of the net size of two generalized distortions, the values of which would depend on the size of the policy variable under consideration.

Frank Alan Ward
Economics Department
Colorado State University
Fort Collins, Colorado 80523
Fall, 1977

ACKNOWLEDGEMENTS

The author gratefully acknowledges the assistance of his committee members, John R. McKean, S. Lee Gray, Dwight Blood, and Robert Gaines, in their help in seeing this dissertation to its completion.

Professor S. Lee Gray, in addition to helping the author express his views in this dissertation, has been invaluable in helping the author improve his economic articulation on other subjects, and in helping the author to gain confidence in himself as an economist. For this, the author is grateful to Professor Gray.

Professor Dwight Blood has helped the author keep this dissertation, and economic theory in general, in proper perspective. Professor Blood is especially thanked for his assistance in helping the author gain confidence as a teacher of economics.

The author also gratefully acknowledges the assistance of Professor Robert Gaines as the first person who helped the author gain a lasting appreciation for mathematics. Professor Gaines also took much time out of his very busy schedule in helping the author with his mathematics in this dissertation. Although the author accepts sole responsibility for any errors or omissions which may remain in this dissertation, he is especially grateful to Professor Gaines for his time and effort.

The author wishes to extend special thanks and appreciation to his adviser, mentor, and personal friend, Professor John R McKean. Had it not been for Professor McKean, the author would have never taken an interest in economics, let alone gone to graduate school and

and studied it at an advanced level. The author can only recount the innumerable late night hours that Professor McKean spent intellectualizing with and lending moral support to the author. What little intellectual maturity the author has, he owes it in large part to Professor McKean. The author is indeed grateful to Professor McKean for his time, energy, patience, and understanding.

Finally, the author wishes to extend extra thanks to Miss Carol Valleroy, who has lent her unselfish moral support, understanding, and patience to the author throughout the year during which this dissertation was written. It is largely due to her that this work was completed.

TABLE OF CONTENTS

page

LIST OF FIGURES

CHAPTER I: INTRODUCTION TO DISSERTATION
SOME FUNDAMENTALS OF EXHAUSTION THEORY
AND LITERATURE REVIEW

A. Introduction to Dissertation

What is the Purpose of this Dissertation?

The fear of the economic consequences of running out of exhaustible resources is by no means recent. In 1865, William Jevons[1] expressed concern about the ultimately increasing costs of coal extraction, especially how Britain's economic growth could be eventually constrained by the tendency for rising coal prices to destroy her competitive advantage in the world market for manufactured goods.

Technological advance notwithstanding, the growing scarcities of exhaustible resources have continued to alarm modern society, and economists have devoted increased attention to the pricing and efficient allocation of exhaustible natural resources. This concern for an efficient allocation has resulted in a resurgence of interest in the pure theory of exhaustion, at both the level of the individual firm and at the level of the socially efficient allocation across an industry.

The need for an understanding of the pure theory of exhaustion is due, in part, to two specific problems which must be faced and dealt with by economists.

First, there is a need for assessing the desirability of a market allocation of an exhaustible resource. Before the assessment can be made, it must be understood how a free market allocates resources toward the production of an exhaustible resource. Specifically economists should have some sort of an idea how production costs, total known supply, and costs of substitute technologies affect output, price, and the life of an exhaustible resource. The reason that knowledge of this process is important is that once determined, such market allocative responses can be compared to some appropriately defined efficiency norm. Relative to an efficient allocation[2] it is important to know if the market produces an exhaustible resource too quickly, too slowly, or at about the right rate.

The second problem which needs to be addressed is related to the first. Economists can perform more meaningful analysis if it is possible to measure just how well a market's allocation approximates an "efficient" allocation of an exhaustible resource. That is, there is a need to develop a precise measure of the degree to which the actual allocation approximates the ideal.

The purpose of this dissertation is to provide a conceptual means of analyzing these two problems.

What Is Included in this Dissertation?

In addressing these two efficiency-related problems of exhaustible resource allocation, this dissertation draws from and extends the pure theory of exhaustion[3] at three different levels. The first level of this inquiry is an extension and refinement of the comparative statics of the theory of competitive mining under certainty. Of primary concern are the effect on industry price and output over time of shifts if resource availability and technology. At this level, the welfare implications of a competitive allocation are addressed only peripherally. However, at the second level of inquiry, the optimality properties of such an allocation over time are examined. Specifically, it is shown that a perfectly competitive market allocation of an exhaustible resource is only Pareto efficient under certain restrictive conditions. At the third level of inquiry, the degree to which the market breaks down is the subject of concern. Here, the presumption is that in all liklihood, there will be some impediments to a Pareto efficient free-market allocation of the exhaustible resource. Specifically, an extension of currently accepted welfare-loss theory is developed and made applicable to the exhaustive resources sector. That is, what is measured is the welfare cost of any arbitrary suboptimal time allocation of the resource.

Thus, broadly speaking, this dissertation attempts to answer three questions: First, what will be the allocation over time of an exhaustible resource in a perfectly competitive industry and how does that allocation change with changes in the cost of a substitute technology, the resource's extraction costs, the discount rate, and the supply of the resource? Second, what are the welfare implications of such a competitive allocation? Third, how can one measure the welfare-cost of a suboptimal allocation of the resource and how might one compensate for such a suboptimal allocation?

Previous Work from which this Dissertation Draws.

As is the case with this dissertation, very few theoretical inquiries can lay claim to total originality. In this work, there is heavy dependence on authors in both the fields of exhaustion theory and applied welfare theory. On a chapter-by-chapter basis, the most relevant previous contributions on which this dissertation rests are as follows:

In the second chapter, which deals with the comparative statics of a market allocation, the most important earlier contributors on the subject are Gray, Hotelling, Herfindahl, and Nordhaus.

In 1914, L.C. Gray[4] showed that it is profitable for a competitive mining firm to restrict its output short of that level at which price equals marginal cost. In 1931, Harold Hotelling[5] showed that a competitive mining industry will tend to allocate an exhaustive resource over time in such a way that its firms' marginal profits increase at a constant percentage rate. Herfindahl, in 1955,[6] showed how cost curves could be used to come to Gray's verbal conclusions, and in 1965,[7] he developed a graphical representation of Hotelling's mathematics, and thus made Hotelling's article accessible to a wider range of readers. In 1971, Nordhaus[8] developed the notion of a "backstop" technology of zero (exhaustible) resource costs, a technology which will radically alter the allocation of exhaustible resources.

In chapter three, which deals with the optimality attributes of a free-market allocation, especially important earlier writers are Hotelling, Gordon, Peterson, and Goldsmith.

Hotelling, in his same 1931 paper, developed the notion of an optimal allocation of an exhaustible resource over time and explored the degree to which alternative market structures would lead to that allocation. Gordon, in 1967,[9] developed an argument showing that a perfectly competitive mineral industry would, in general, fail to allocate the exhaustible resource efficiently. Peterson, in 1972,[10] used variational methods[11] to show that Gordon had erred and that a competitive economy would, indeed allocate the exhaustible resource efficiently. Also central to the optimality attributes of a market allocation is the work of Goldsmith, who in 1974,[12] supported Peterson's contention and gave a numerical example disproving Gordon's 1967 assertion.

In the final section of the dissertation dealing with the welfare cost of a resource misallocation, the primary sources from which the theory is extended are those of Harberger (1971),[13] and Boadway (1974).[14] Harberger developed a simple, yet intellectually satisfying measure of welfare loss in a normal flow market setting, while Boadway reconciled the approaches of two different groups of applied welfare theorists by showing that both of their welfare-loss measures were, in fact, equal. In this section of the dissertation, where a welfare loss measure is developed for any arbitrary time allocation of an exhaustible resource, the works of Harberger, Boadway, and others are used in order that an appropriate theoretical basis might be provided for the inquiry.

B. Some Fundamentals of Exhaustion Theory

Exhaustible Resources and Economic Theory

In recent years, with the development of a conservation ethic, there is an increasing suspicion that the free market allocation of exhaustible resources does not completely reflect the needs of future generations. However, economists realize that the issue between the growth and antigrowth contingent is:

> "not whether and how much provision must be made for future generations, but in what form it should be made. The conservationist emphasizes exhaustible resources--minerals in the ground, open space, virgin land. . ."[15]

Economists also point out that:

> "In a properly functioning market economy, resources will be exploited at such a pace that their rate of relative price appreciation is competitive with rates of return on other kinds of capital. Many conservationists have noted such price appreciation with horror, but if the prices of these resources accurately reflect the scarcities of the future, they must rise in order to prevent too rapid exploitation. Natural resources should grow in relative scarcity--otherwise they are an inefficient way for society to hold and transmit wealth compared to productive physical and human capital."[16]

Solow explains the capital asset nature of an exhaustible resource very eloquently, and we draw heavily from him.[17]

A pool of oil or vein of iron or deposit of copper in the ground is a capital asset to society and to its owner much like a printing press or a building or any other reproducible capital asset. The only difference is that the natural resource is not reproducible, so the size of the existing stock can never increase through time. It can only decrease. This is true even of recyclable materials. A formula just like the ordinary multiplier formula tells us how much copper use can be built on the world's initial endowment of copper, in terms of the recovery ratio. There is always less ultimate copper use left than there was last year, less by the amount dissipated beyond recovery during the year. So copper remains an exhaustible resource, despite the possibility of partial recycling.

A resource deposit can only produce a current return for its owner by appreciating in value. Capital asset markets can be in equilibrium only when all assets in a given risk class earn the same rate of return, partly as capital gain. The common rate of return is the interest rate for that risk class. Since resource deposits have the property that they yield no dividend so long as they stay in the ground, in equilibrium the value of a resource deposit must be growing at a rate equal to the rate of interest. Since the value of a deposit is also the present value of future sales from it, after deduction of extraction costs, resource owners who are not selling must expect their marginal profits to be increasing exponentially at a rate equal to the rate of interest. If mineowners' marginal profits are increasing at the rate of compound interest, they will be indifferent between receiving marginal profits now equal to MP_0 or marginal profits equal to MP_0e^{rt} at any point "t" in the future.

What are the market mechanics which cause marginal profits to rise exponentially? Firms realize that due to their fixed mineral supplies, it is privately profitable to restrict current output below that level at which price equals marginal private cost. They realize that exhaustion imposes real costs on them, and therefore they are forced to take account of this depletion "user cost," the cost of using the mineral now, rather than at its best alternative time in the future.[18] Thus we should expect rational mining firms which have a fixed resource deposit to adjust their time path of output so that marginal profits, rather than being zero, have an equal discounted value in all periods.[19] That is, since increased current output decreases the amount that may be sold at a future period, if the marginal profit on current output has a lower present value than that on future output, a firm will restrict current output and sacrifice current profits, thus obtaining more valuable future output. Once firms adopt this output pattern, they have done the best they can to minimize the ultimate cost of exhaustion, i.e. they have minimized their user cost.

Firm behavior is easily generalized to industry behavior, as explained by Solow. Firms react to any given price path over time in such a way that marginal discounted profits are equal in all periods; that is, for any arbitrary price path, each firm chooses a well-defined output path. The determination of industry price comes, of course, from the output of the firms and the industry demand curve, which in turn, leads to market-clearing at all points in the future. So, ultimately, when the market price rises, the current rate of production must fall along the demand curve. Eventually, market price is so high that it chokes off demand entirely. At that moment, production falls to zero as the last mine has produced its final remaining ounce of ore. If the exhaustible resource tracks its price-output path well, the last ounce sold will be the last ounce in the ground, and the resource will have been exhausted at the instant that it has priced itself out of the market.

Optimality Properties of a Market Allocation

As in the case of other competitive equilibrium models, the competitive model of an exhaustible resource has certain optimality properties, provided that certain assumptions are met. Given the appropriate assumptions, as was pointed out by Harold Hotelling in 1931[20] the resulting industry equilibrium maximizes the sum of discounted consumer-plus-producer surplus, provided that society wishes to discount future surpluses at the same rate that mine owners wish to discount profits; If this is so, it can be said that the allocation of resources toward the production of the exhaustible resource is Pareto efficient.[21]

Without attempting a review of the literature on the Pareto efficiency of competitive equilibrium, it is safely said that there is general agreement that the conditions under which an exhaustible resource is efficiently allocated are the same as those for a renewable resource, with a few qualifications and extensions.

Some Sources of Market Failure

First, it is generally thought that a set of well-functioning futures markets and contingency markets are needed to ensure an efficient allocation, as time plays an especially crucial role in the exhaustible resources market. Futures markets are important because the entire concern about exhaustible resources is that of allocating a fixed stock between competing uses at alternative times rather than between uses at a single point in time. With no futures markets, resource traders will have to base their intertemporal allocative decisions on their guesses of future prices, which may not be at all the same as equilibrium future prices. These guesses, of course, influence planned intertemporal output, including current output and thus current prices. In this situation, it is surely reasonable to ask if current prices are really equilibrium prices, or if they are so unstable as to be allocatively wasteful and socially detrimental. Contingency markets that insure against such eventualities as the failure of predicted technologies, are necessary to deal with uncertainty. In any economic decision for which events in the distant future play an important role, great uncertainty will generally lead to market failure unless there are markets for contingent commodities.

In addition to this, the discount rate probably plays a more crucial role in the time allocation of exhaustible resources than in that of renewable resources. In the market for exhaustible resources, not only does the interest rate determine the allocation of capital, as it does in other markets, but it influences the allocation of the resource itself. Solow points out that with exhaustible resources added to the picture, the optimal path of output (in a one-sector economy) with a positive discount rate calls for per-capita consumption to go to zero, whereas with a zero discount rate, that path would call for continually rising consumption per head. That is, even when technology and available resources allow for a rising standard of living, positive social time preference could lead society to prefer eventual extinction, arising from the exhaustibility of natural resources. If it is true that the market interest rate exceeds the true marginal rate of time preference, then the price of the resource will rise faster than it should, as discounted marginal profits are worth less to producers than they "should be." Thus, current production will be carried too far, and since the resource supply is fixed, there is not enough production in the future and the resource will become exhausted too soon.

The above sources of intertemporal market failure, lack of futures or contingency markets, and interest rate distortions are particularly applicable to exhaustible resources. Other sources of market failure, also common to the renewable resources markets, will be mentioned briefly.

Monopoly, by restricting output, underallocates production in flow markets. Monopoly is also generally excessively restrictive in resource markets. Whereas the competitor equates price-less-marginal cost to user cost, the monopolist equates marginal revenue-less-marginal cost in all periods. This typically prolongs the period of extraction beyond that which is socially desirable.

Common property ownership presents a well-known source of market failure and is especially detrimental to an efficient allocation of the exhaustible resource. If, by withholding a unit of output from today's market, a producer can extract that unit for sale at a later date, a socially efficient allocation may result. If, however, by not producing that unit today, he knows that his competitors, who are co-owners of the resource pool, will take it from him, a cost is imposed on the producer which is not a true opportunity cost for society. As is the case in normal flow markets, output is higher (in earlier periods) than is socially warranted.

Attempts to deal with the common property problem, expecially in the case of oil, have led to another type of market failure. Since governments realize that unregulated common pools lead to wasteful extraction, prorationing laws (limiting the output flow of some producers) have been initiated. Prorationing, although potentially less wasteful from the standpoint of engineering efficiency, is arbitrary, and fails to allocate the resource economically efficiently. In principle, the ideal solution is to place the pool's allocation decision under a central management, whose goal is to maximize joint extractors' profits.

There are other sources of market failure in the exhaustive resources sector, but since most of these sources are not particularly unique to this sector, we will touch upon them only in passing when discussing specific contributions to the literature.

C. Review of the Literature on the Theory of Exhaustible Resources

The literature dealing with the theory of exhaustible resources typically addresses one or more of the four following very broad issues: First, there is a need for an explanation of how various market structures do allocate the resource. Normally, such an explanation is founded on behavioral assumptions about both the individual firms and the industry.

Second there is needed an explanation of the attributes of a desirable allocation, including that of the exhaustible resource itself, resources related to its extraction, and resources devoted to the development of a substitute technology. The attributes of that desirable allocation of course, depend on the specification of a social objective function. The social objective function is not necessarily unique. It could, for example, relate to efficiency in either a micro setting, e.g. surplus maximization, or a macro setting, e.g. a measure of time-discounted utility of consumed aggregate income. In fact, the objective function could even relate to some measure of equity, either between present actors, or across generations.

The third issue having to be addressed is that of the degree to which one and two approximate each other. That is, how close is the behaviorally-determined allocation to the ideal allocation? The measure can be either qualitative (either the two are or are not "close" to one another) or quantitative (a measure of how far apart they are).

Finally, there has been some broad concern in the literature for means of compensating for market failures. Such studies have often addressed the issue of compensatory taxation, either for efficiency or equity purposes, but have also sometimes concerned themselves with the means of correcting for a divergence between private and social rates of discount.

With this discussion of (1) how the exhaustible resource sector fits into the overall scheme of an economy's allocation of resources and (2) the issues to which most of the theoretical literature has directed itself, let us outline some of the major contributions to resource exhaustion theory as of this date.

In 1914, Lewis Gray developed what is probably considered the first attempt at a unified theory of the mine under certainty.[22] In incorporating the theory of Ricardian rent as applied to exhaustible assets, he made the first major distinction between the rational operation of a mine and that of the typical profit-maximizing firm. Gray noted that although there is a tendency for a mine owner to stop short of the point where price equals marginal cost, a positive discount rate keeps him from restricting his output to the level where average cost is a minimum. His second contribution to the theory was his verbalization of what Hotelling was to show seventeen years later in 1931[23]--that the mineowner adjusts his output rate so that the present value of marginal profits in current and future uses are just in balance, even if the extraction cost function shifts over time due to cumulative extraction. He also hinted at a point brought out by Gordon in 1967,[24] namely that with increases in expected future price, relatively more output is allocated to later periods and less to earlier periods.

Gray also set the stage for many later papers in his discussion of the allocative effects of taxation. Some of his observations are as follows: An annual tax on the value of the mine encourages present production as opposed to future production, and will probably lower present price. A tax upon annual economic profit does not affect the firm's time path of output, and a tax which is perceived as temporary will have the effect of putting production off into the future.

Gray showed some interest in the theory of optimal capital accumulation as applied to the extraction of an exhaustible resource, a subject which was not fully generalized until 1970 by Burt and Cummings.[25] It was Gray who recognized that to the extent that a continual decline in output over time means wasted capital that is only used in the earlier periods of production, a mineowner will attempt to make output more nearly uniform throughout the life of the mine.

Finally, and probably his main reason for writing the article, Gray faced the question of how much depletion charge should be deducted from the cash inflow of a mine. Gray's concern here was on the allocative significance of profits over and above that required to maintain the mineowner's capital. His conclusion was that net royalty, since it is not a part of supply price, can be taxed away and not affect output, provided that the relation between present and future is not upset.

In 1931, Harold Hotelling published what is considered the most significant contribution to the theory of exhaustible resources.[26] Influenced by Ramsey's pathbreaking work using the calculus of variations,[27] Hotelling applied Ramsey's mathematics to the dynamic problem of the intertemporal allocation of an exhaustible resource.

He was the first to prove convincingly that in order that maximum social value be achieved (consumer-plus-producer surplus), unit profits must grow exponentially.

However, Hotelling's results, impeccable as they were, did not generalize far enough. Whereas Hotelling inferred that under an ideally competitive system, the exhaustible resource will be allocated efficiently if marginal profits grow exponentially, he did not address the issue of end-point conditions. This issue has been pointed out by recent authors and is discussed in this dissertation.

Hotelling was the first author to point out that a monopolist, rather than exploiting the resource too quickly for his own gain, will in all likelihood, restrict output too much for the social good. It has been pointed out since then that the monopolist and the conservationist are probably in more agreement than either thinks.

Hotelling also extended Gray's taxation analysis. He showed that an anticipated taxation of the capital value of the mine of "A" per cent per year has the same effect as increasing the private discount rate by "A" per cent; further he showed that severance taxes tend to conservation. Although he did not explore it in great depth, Hotelling implied that in general, there exists a path of tax over time, a function of the rate of production, cumulative production, and time, which can lead to a maximum social value of the resource.

In 1944,[28] S.V. Ciriacy-Wantrup also examined the allocative effect of several kinds of taxation. He showed that progression in taxation leads to increased conservation, present-value taxes tend toward depletion, while income taxes are more nearly neutral. Wantrup concluded that state and local governments should make greater efforts at income taxes and depend less on taxes which tax the value of the mine.

Anthony Scott published a paper exploring the economic theory of user cost, in 1953.[29] Although not directed specifically at the theory of exhaustion, Scott's contribution remains an important addition to the theory of exhaustible resources. Scott defined user cost as the decreased value of capital assets (including depletable resources) which results from an increment to current output. User cost generally rises with higher output, but as in the case of a forest that needs thinning, need not necessarily.

At a higher level of abstraction, Scott argued that user cost could be viewed as the future opportunity lost due to another economic decision being carried through. Scott's article on user cost is an important part of the basis of the current notion that exhaustion of a depletable resource can be treated within the broader framework of a capital asset depreciation.

In 1954, Donald Carlisle brought the theory of exhaustible resources down to the level of increased practicality, from his experience as a geologist.[30] Carlisle was probably the first to point out to economists that not only is the rate of mineral output a decision variable, but the total cumulative quantity of ore is often a variable subject to choice. Economists until this time had generally ignored the total quantity extracted as being dependent on the time path of the extraction rate. Thus, for example, Carlisle pointed out that, for a given capital stock, if a very rapid rate of extraction is deemed necessary (i.e. a rapid rate of extracting the <u>mineral</u>, and not just a rapid rate of <u>ore</u> extraction), a very selective type of mining will be chosen, which ultimately leaves a greater amount of the mineral in the ground. Carlisle, then provided a conceptual model of his own whereby the mineowner maximizes the value of the mine by choosing values for both variables.

Carlisle also drew attention to some of the practical difficulties involved in extraction of the mineral and in capital accumulation. Specifically, he noted that uncertainty combined with institutional practices often lead to building a smaller mine than would otherwise seem appropriate. First, mineowners are rarely, if ever, aware of the exact size of the deposit. For this reason, mineowners optimize their time path of mine construction by expanding mine capacity hand-in-hand with mineral discoveries. With the smaller mine capacity, mineowners produce at a lesser-than-theoretically-optimal extraction rate, in an attempt to prolong the life of the mine for two reasons. First, since total reserve quantities are usually unknown, the hope of finding more reserves is one of the strongest incentives for expanding the life of the mine. In addition, the life of the mine may be expanded because of the desire to attract a more permanent class of worker, or to meet a smelter's need for steadiness in mineral input.

Carlisle also pointed out that risks of the mineral market enhanced the possibility of mining too selectively for the social good. Thus, although not alluding to it by that name, Carlisle, in effect, brought attention to these risks as a likely source of <u>market failure</u>.

Orris Herfindahl, in 1955, emphasized that there are economic forces behind the exploration process, and therefore drew attention to exploration as being an intimate part of the exhaustion process.[31]

Among Herfindahl's contributions include a proof that with constant exploration costs, the mineral sector is no different than any other sector and there will be <u>zero</u> long-run royalties[32] occurring in the industry. Furthermore, Herfindahl showed that given increasing <u>cumulative</u> discovery costs, the stock of properties will decrease with further rises in price, until further discoveries are impossible, after which unit royalties will rise at the rate of compound interest. In addition, if further discovery is possible only with rapidly rising exploration costs, the bulk of current supply will come from reduced stock, and although price will not rise at an exponential, exploration will not be in sufficient volume to <u>maintain</u> the stock of properties.

In 1967, a book which was compiled of papers by several economists, grew out of a 1964 conference at the University of Wisconsin on "Tax Treatment of Exhaustible Resources," sponsored by the Committee on Taxation, Resources, and Economic Development.[33] The resulting book contains several papers that are pertinent to the theory of exhaustion.

Orris Herfindahl presented a paper which developed the simple geometrical comparative statics of exhaustion theory.[34] In the paper, Herfindahl explained the sensitivity of price, royalties, and exhaustion time to the following: changes in costs of one of several grades; changes in quantities of several grades; changes in the discount rate; and changes in the elasticity of substitution in production. This dissertation's chapter two is, in part, an extension of Herfindahl's paper.

In his paper, Herfindahl advanced an explanation as to why prices do not seem to rise as the theory of the mine would suggest; namely, it is because of technological advance over time and the continual discovery of new and richer sources of supply. Evidently, Herfindahl felt that those discoveries yielded unexpectedly rich sources, for his 1955 paper would have incorporated discoveries as part of the exhaustion process.

In the same volume, Anthony Scott developed what could reasonably by considered a synthesis of the theory of the mine as of that date.[35] At the level of social efficiency, Scott pointed out to us that is privately and socially economical to use up our best opportunities first, as opposed to the natural reaction of a conservationist who thinks that we should preserve our best deposits for a longer time. Scott also generalized Gray's nation that the socially optimal rate of output is a balance between time preference and increasing marginal costs associated with the higher rate of output.

Scott also generalized the theory of the individual mining firm by showing that the mine owner tilts his production plan in favor of the present, rather than producing at constant rates. This is not necessarily because of the deterioration of ore quality, falling prices, or even due to costs rising over time. The tilt can exist simply due to rising marginal costs as a function of the output flow. What Hotelling showed for the industry, Scott showed for the individual firm.

In addition, Scott extended Carlisle's discussion on the factors entering into the determination of the appropriate cut-off grade, added to Wantrup's analysis of the effects of taxation on mineral output, and with the aid of some elementary diagrams, attempted a generalized theory of the optimal size of the mine under conditions of certainty.

Also contained in the same volume was an article by William Vickery, part of which dealt with potential sources of market failure.[36] Among other subjects, Vickery noted that although competition tends to be Pareto optimal, imperfect possession (common property) and imperfect foresight, which causes social and private risks to diverge, can lead to excessively rapid depletion.

In this volume, a paper by Stephen McDonald developed the social optimality conditions with respect to exploration rates.[37] According to McDonald, that rate at which unknown reserves are transformed into known ones must be such that both the marginal yield on postponing the use of an increment of the known stock, and the marginal yield on exploration for renewal should be equal to the marginal rate of time preference adjusted appropriately for risk and uncertainty. McDonald could have significantly added to the paper had he explored the tendency to which the free market achieved this allocation.

In the editor's conclusion, Gaffney spent some time showing where the exhaustive resources sector was likely to suffer a market failure, especially due to institutional restrictions.[38] Specifically, he pointed out that oftentimes, private capital is wasted on the development of submarginal reserves due to an archaic conservation ethic, while prorate schemes also limit the efficiently rapid use of superior deposits, especially in the case of oil. In a similar light, he argued that obstacles to survey in superior regions should be eliminated, for resources are wasted by diverting exploration to inferior deposits.

Gaffney also strongly urged taxation on principles of equity. Specifically, the public should place heavier taxes on superior resource deposits, and in general, economic rent on minerals should be more steeply taxed, the revenues being used to stimulate public exploration, as private exploration is redundant and wasteful.

In 1966 and 1967, Richard L. Gordon published two innovative, refreshing, and somewhat controversial papers.[39,40] In his 1966 paper, Gordon took issue with the commonly-accepted theoretical notion that an upward biased interest rate leads to excessively rapid depletion. In that paper, Gordon argued that high interest rates also increase capital costs, which have the extra effect of <u>lessening</u> capital intensity and making some otherwise profitable deposits prohibitively expensive. The effect would be a lower overall rate of current extraction. Gordon constructed a simplified model employing assumptions of a capital theory of cost where numbers were generated showing that higher interest rates <u>can indeed</u> prolong the exhaustion period. The main policy implication of this paper is the following: If there is any upward bias to the private rate of discount, a reduction in this divergence may increase rather than reduce the use of exhaustive resources.

Gordon's 1967 paper is probably the first to have as a main goal the comparison of an optimal central management of an exhaustible resource and that same resource mined by a competitive industry. That is, in this paper, Gordon explicitly addresses the question of the Pareto optimality of a competitive market allocation. In this paper, he brought together the theory of the individual mine, combined it into an analysis of the industry, and compared the allocative efficiency of the outcome with that of a hypothetical efficiency planner.

Gordon also noted some rather peculiar implications of the pure theory of exhaustion. First, since exhaustion theory suggests that it is profitable for mines to sacrifice some current profits for future profits, and since Gordon does not believe that mines are generally seen restricting current output below the profit-maximizing level, the theory would seem to be telling us that either conservationists are concerned about a non-existent problem, or that the effects of exhaustion are so small that present generations are wise to ignore it. Another implication of the theory which was noted by Gordon is that one is led to believe that exploration will be undertaken only if the present value to the loss to excessive output-reduction exceeds exploration costs. However, since he does not see mines restricting current output below their short-run maximizing levels, perhaps potential losses, and hence the value of better reserves are zero. Therefore, Gordon would seem to argue that properties are explored only because firms want to exploit them.[41]

Gordon's 1967 paper was generally a very well-thought-out contribution to the literature. However Gordon made one rather major error when he stated that the purely competitive market exhausts the resource too quickly. Whereas he believed that efficiency requires that production costs in a given period should be minimized (i.e. equate marginal production costs across all mines), in fact it has been shown since then[42] and is also shown in this dissertation that since exhaustion is a real social cost, efficiency requires that marginal social costs, including the cost of exhaustion (user cost), is that which should be equalized across mines.

In 1968, Ronald Cummings wrote a Ph.D. dissertation and later published part of it as a journal article.[43,44] Both dealt with some generalizations of the pure theory of exhaustion, using the new mathematics of optimal control theory.

In the works, Cummings generalized the theory as follows: By incorporating the effect of cumulative extraction on costs, he proved that the effect of cumulative production leading to increased costs is to (1) lower the rate at which marginal profits grow, and to (2) stretch out the period of exploitation, the latter being a result suggested by Scott in 1967.

Cummings also developed optimality conditions for a single mine which had secure tenure in its resource deposit. He finally determined the optimal production path forthcoming under a central authority which directs the output of several common-property mines. Here, although Cummings claimed that his model showed "optimal paths for several common property firms desiring to maximize joint profits" was a condition for social efficiency, this is questionable, since surplus measures were not specified in his objective function. Perhaps Cummings meant "efficiency for the industry," rather than for society as a whole. In any case, the papers would have been strengthened had they explored the social allocative effects of his various regimes.

In 1968, Vernon Smith constructed a generalized model describing the dynamic process of recovery of several kinds of natural resources.[45] The paper constructed the equations of motion of both capital (number of firms) and industry output, based on a profit-loss incentive system. The paper was global in scope, in that it explored the effect of economic motives on the equilibrium of biological systems. The main weakness of the paper, from an exhaustible resources standpoint, was that it neglected the entire tilt of current production forthcoming from firms which realize that exhaustion imposes real costs on them. Therefore, as the mechanism for firm entry-exit was simply current profit flows, the paper really contributed more to mathematical biology than to the theory of exhaustible resources.

Oscar Burt and Ronald Cummings, in 1969[46] and 1970,[47] coauthored two papers which substantially generalized the theory of capital accumulation as applied to the theory of the mine. In their 1969 paper, the authors built upon Smith's observation that there is an intimate relationship between capital and resource extraction; however their model was an advance over Smith's in that it also conformed to the established theory of resource exhaustion. Specifically, the 1969 paper developed, for a single firm, simultaneous optimality conditions for capital accumulation and rates of production of the resource. However, their 1969 paper was only a starting point for their 1970 contribution. The latter was a highly generalized and comprehensive model for simultaneoulsy optimizing the rate of resource extraction <u>and</u> investment in resource industries.

In their 1970 paper, using the discrete maximum principle, the authors stated the objective function, production relationships, and investment functions generally enough so that the functional forms relating to any given resource could be plugged into the model and the optimal solution would fall out as a special case. For example, the objective function was generalized to the extent that optimal time paths of extraction and investment could be determined, from the perspective of a firm, an industry, or society-at-large. A particularly useful optimality property of an investment-extraction system was pointed out by the authors and is as follows:

> "At all points in time, the marginal social value of current production should be equal to the discounted marginal value of a unit of resource retained in stocks rather than in current production plus the discounted marginal value of capital stocks consumed by an increment to current production (the second two terms being the total real user cost)."[48]

In 1972, Frederick Peterson, using the calculus of variations, extended Hotelling's original work, and pointed out the error in Gordon's 1967 paper.[49] Among other results, Peterson's work showed that in the absence of market imperfections, a competitive mining industry extracts an exhaustible resource in such a way as to maximize discounted consumer-plus-producer surplus, a result contrary to Gordon's 1967 assertion.

Peterson also included provisions for technical progress, exploration, and recycling, under conditions of both a fixed stock of a resource and a stock that is limited only by rising costs associated with depletion.

A major contribution to the literature was Peterson's theorem that described the set of taxes that do not alter firm behavior over time. This was a generalization of previous authors who had attempted to describe the allocative effects of several types of taxation.

Neil Vousdon, in 1973, developed a theoretical resource depletion model which, using the principles of neoclassical growth theory, explored the allocation of an exhaustible resource which maximized intertemporal utility.[50] This model was neither a firm nor an industry model; rather Vousden's one-sector model assumed that aggregate production depends on the inputs of an exhaustible and an inexhaustible resource. Using optimal control theory, Vousden, by adding the twist of a conservation motive in society's utility function, showed that it may be optimal to never exhaust the resource. He also showed that if institutional constraints prevent the extraction rate from falling below a given finite rate, then it is never optimal to avoid exhaustion in a finite time. However conservation motives can be taken into account by lowering the extraction rate to its lower bound earlier in the program.

A major contribution of Vousden's paper would have to be that it provided theoretical sanction to the notion that given a strong conservation motive,[51] where the resource in its natural state is valued in addition to its value as an input, there is justification for permanently postponing its total extraction.

In 1973, William Nordhaus developed a unique model that investigated the efficient allocation of energy resources over time.[52] In the model, Nordhaus is probably the first author to actually attempt a calculation of a set of efficient price paths for a set of various energy resources, using empirical data. By assuming that a "backstop technology" would be available within 200 years (a technology, although expensive relative to today's standards, one which will rest on an effectively infinite resource base), Nordhaus traced the efficient price path of current resources--a path which provides the cheapest way of meeting a growth path of final demands for energy products. Using some conservative figures about the given stock of energy resources and a given set of processes for converting resources into products, Nordhaus came up with some rather surprising results. He concluded that his optimal prices were close to actual current resource prices with the exception of petroleum; petroleum's current price was about 240 percent of the efficient price calculated in the optimal program. With this result, Nordhaus noted that as a long-run policy, it would be unwise to arbitrarily increase the price of energy products in the interest of artificially preserving energy resources. Especially, further increased in the price of gasoline (via taxation) would be unwise.

In 1974, F.E. Banks published a paper which was essentially a critique of the 1973 Vousden paper.[53] According to Banks, Vousden's paper consisted of merely repeating what was already known about exhaustion theory, but using the technique of optimal control theory. Further Banks believed that Vousden's conclusions were of questionable relevance, as Vousden used an intertemporal utility function in his objective function, rather than a social-private profits function, which Banks believed would have had more economic content. Perhaps Banks was overly critical of Vousden's paper, as it would appear that Vousden's major contribution lay in his incorporation of the conservation motive into the social welfare function, showing that the resulting allocation would call for leaving some of the resource in the ground.

In 1974, the Review of Economic Studies published the papers of several authors who participated in a British natural resources symposium, from which we now discuss three particularly relevant contributions.

The first paper, by Milton Weinstein and Richard Zeckhauser, is a generalization of the optimality properties of a competitive market allocation--to the case when interest rates are determined endogenously, i.e. by the extraction process itself.[54] The paper also incorporates the efficiency attributes of a competitive market which recycles the resource at a positive cost. The paper proved that, in general, neither of the two factors detracted from the overall social efficiency of competition.

The second paper to which we make reference is that of R.M. Solow. Solow's paper dealt with the implications of intergenerational equity in an economy where resources are exhaustible, an issue which has never been fully resolved in the literature.[55] Probably Solow's major quantifiable result was a proof that, in a neoclassical growth model, if the elasticity of substitution between the exhaustible resource and neoclassical factors is greater than unity, a constant level of consumption can be maintained indefinitely.[56] This is a particularly useful result, as it would seem to satisfy even the most conservative criterion of intergenerational equity.

The third paper, by Partha Dasgupta and Geoffrey Heal, agreed with and extended Solow's results.[57] Dasgupta and Heal, also dealing with a neoclassical one-sector economy, showed that in an economy with both reproducible capital and an exhaustible resource, an optimal program should have the capital-resource ratio changing at a percentage rate equal to the product of the elasticity of substitution and the average product per unit of fixed capital. Thus, the easier it is to substitute, and the more important is the reproducible input, the more necessary it is to be able to substitute the reproducible resource for the exhaustible one.

It should be remarked that both of the "one sector" papers modeled the exhaustible resource within the confines of a neoclassical one-sector economy, dealing with aggregate output rather than addressing the goal of efficiency. Therefore, their results should be interpreted tentatively.

In 1974, R.M. Solow published another paper on exhaustible resources, this one as the Ely lecture at the 1973 American Economic Association meetings.[58] Although primarily designed as an introduction for economists to the fundamentals of exhaustion theory, Solow made several worthy comments.

A major emphasis of the paper was on the stability of resource markets. Whereas most past observers had emphasized the instability of price expectations leading to wide swings in price and output, Solow noted that the capital asset nature of the market would lead to more stability than previously thought. According to him, when the current resource price is low, rather than resource owners hastening to add to current production, they will assume a capital loss by writing down the value of the ore, at which time, production proceeds at a stable rate over time. Solow's argument, therefore, gives some hope for the prospect of exhaustible resource markets tracking their equilibrium paths well.

Although more work needs to be done in the area of expectations and stability in the absence of futures markets, Solow gives economists a bases for at least cautious optimism. Stability issues aside, however, Solow made the rather sobering observation that the degree of substitutability between exhaustive and reproducible resources together with the possibilities for technological advance are the two major factors which will ultimately determine society's ability to withstand resource shortages.

Ngo Van Long, in 1974, wrote a short note on resource exhaustion theory, which would have to be considered a technical note, although worthy in that respect.[59] Using a simplified diagrammatical approach, Van Long emphasized that user cost, the maximum marginal value of a unit of the resource at any other time period rather than the present, like many other shadow prices in optimizing models, comes out of the solution of the problem rather than being of any help to obtain the solution. Long's paper is valuable particularly as it makes the essence of the theory of the mine readable to a wide variety of readers.

Also in 1974, Oliver Goldsmith published what was primarily a rebuttal to Gordon's 1967 paper.[60] Using the calculus of variations with constraints, as did Peterson (1973), Goldsmith argued against Gordon's contention that a perfectly competitive market will fail to achieve an intertemporally efficient solution.

Goldsmith pointed out, in his rebuttal, that Gordon erred in thinking that an efficiency planner will minimize production costs in a given period by equating marginal production costs among operating firms. Rather, he contends that should the efficiency planner choose this course of action, although <u>current</u> costs would be lowered, <u>future</u> costs would be raised too much, since the most efficient mines become depleted too soon, whereby production would have to be increased from the higher-cost mines. Goldsmith concludes that equalization of marginal costs across mines is an efficiency condition only in the sense of cost defined as the sum of direct production <u>plus</u> user cost.

Goldsmith's proof would have been more convincing, however, had he shown that both regimes (competitive and efficiency) not only faced identical Euler equations, but identical transversality conditions. That is, to prove what he wished to prove, he should have appealed to the necessary "end point" conditions in the calculus of variations, those conditions which establish the optimal output in the final period of extraction.[61]

In 1974, Vernon Smith published a paper combining the theory of exhaustible resources with two-sector growth theory and the Nordhaus conception of a backstop technology.[62] In keeping with the style of the day, optimal control theory was employed. Smith's approach was to maximize the time-discounted value of aggregate output, producible from either a low-cost exhaustible resource in the ground or a high cost inexhaustible resource.

According to his results, along the optimal path, the economy first specializes in recovery of the exhaustible material from the earth, whereby as the extraction rate of the resource declines due to a declining marginal productivity of labor, the backstop technology becomes competitive. Over time, labor begins to flow into the backstop sector at the same rate as the resource is depleted. Ultimately, labor is totally employed in the backstop sector.

Smith's paper concluded on an optimistic note; however one wonders why he did not compare his idealized outcome with a competitive market outcome. Such a comparison could have generated substantially more ground for optimism.

In 1975, The Economics of Natural Resource Depletion was published, arising from a series of papers presented at a conference on resource depletion, organized by the United Kingdom Environmental

Economics Study Group and the Institute of Environmental Sciences.[63] From this book, three papers of particular relevance are mentioned:

A paper by Surrey and Page, written in a casual style, spent some time criticizing the "doomsday models" of the day in addition to advancing a case for increased government subsidization of substitute technologies.[64] In essence, the authors argued that the market fails to allocate enough resources to research and development for substitute technologies. They believe that the market fails because of inherently long lead times (and presumably, because of risk-averse entrepreneurs) and high uncertainty as to future mineral supplies.

In the same volume, Geoffrey Heal outlined the consequences for the allocation of depletable resources given the non-existence of forward markets and contingent commodity (risk) markets.[65] Forward markets are especially important in the case of exhaustible resources, for with such resources, a proper allocation is needed between current and future uses, whereas with renewable resources, one's prime concern is the allocation between competing uses at a point in time. Heal notes that in the absence of futures markets, future prices are <u>expected</u> prices, and the market's failure to predict accurately will lead to wide price and output fluctuations. In fact, Heal tentatively expressed the belief that this is a reason for the empirical observation of wide fluctuations in actual resource markets.

Contingent commodity markets are important because whenever events must be accounted for in the distant future, uncertainty plays a major role, and especially because resource markets are characterized by great political risk. Given this high uncertainty, risk-averse resource traders will deplete the resource too rapidly for the good of society. However, Heal notes that even with no such markets, the appropriate depletion rate might come about if there is some combination of risk-averse and risk-neutral individuals. Heal also notes that efficiency might result if the time distribution of utilities dictates corner solutions.[66]

The third paper in the book to which reference is made is one by John Kay and James Mirrlees.[67] These authors attempted to ascertain whether exhaustible resources are, on a net basis, being depleted too slowly or too rapidly, relative to the competitive outcome. They begin their analysis by exploring the stability properties of the system. Here, agreeing with Solow,[68] the authors feel that there is a good chance for the market to track its equilibrium path reasonably well, as long as price anticipations are governed by past as well as present prices in a positive way--and the assumption of interpolative expectations in the long run is reasonable. Thus, they tentatively place faith in the stability of equilibrium; however, they do advance several reasons for a belief that the market might very will be using up exhaustible resources too slowly.

Kay and Mirrlees note that often times the observed price of exhaustible resources is much higher than the marginal costs of extraction and transportation, and since it does not seem likely that an actual price below the extraction cost could persist for long, if there is a price bias, it must be in the upward direction. That is, price will track a higher-than-efficient path, leading to underexploitation of the resource. As a comment on this particular assertion of the authors, it would seem that they are saying that the resource owners are likely to persistently think that current prices are rising too slowly relative to future prices--for the reader will recall that it is the expected change in prices relative to the current change that determines the time profile of profitability, not simply the current level of prices.

The authors also noted that since successful exploration results have certain public good qualities, risk-averse owners are likely to underallocate resources for exploration. In addition to this, profits taxes, although not affecting the time profile of receipts, will decrease them, yielding another loss-of-incentive-effect on exploration expenditures. Both of these phenomena lead to insufficiently rapid exploration, and hence to too low a depletion rate.

Finally, it is pointed out that since governments often impose revenue-raising taxes on the use of exhaustible resources, the effect is to restrict current output and increase the general degree of monopoly, very likely leading to an excessively slow depletion rate. Here, as a comment, it should be noted that the degree to which monopoly increases depends on which firms bear the greatest burden of the output restrictions. If, as in the case of present prorationing statutes, the restrictions apply only to the larger firms, the degree of monopoly could easily decrease.

The authors did admit that there were two strong factors at work in resource markets which tend to excessively rapid depletion, namely uncertainties of the market and an upward bias of the market interest rate; however the general upshot of their paper is that it is likely that exhaustible resources are being used too slowly, and the interests of future generations would be better served by leaving them "productive equipment rather than minerals in the ground."[69]

In 1975, Dwight Lee and Daniel Orr added to Gordon's 1966 paper by advancing the argument that higher interest rates might decrease the rate of depletion.[70] They emphasized the previously unmentioned role of mineral inventories. Their paper showed that if a mineral can be stored at a finite cost, a monopolist with a wide range of scale economies may lessen his extraction rate with an increase in the interest rate, as higher interest rates add to inventory carrying costs. Thus, whereas the standard "Hotelling" formulation emphasizes that a higher discount rate decreases the value of future profits, itself tending to a greater extraction rate, Lee and Orr pointed out that the rational producer may cut his extraction rate in the face of rising inventory costs. Their argument would tend to reinforce that of Gordon (1966), who argued that the effect of higher interest rates on the cost of non-inventory capital is to discourage the acquisition of machines and hence lower extraction rates. The effect of both of the above papers is to introduce some question into the entire interest-depletion mechanism, pointing to a need for more theoretical work in this area.[71]

In 1975, Weinstein and Zeckhauser extended the conditions under which a competitive market allocates an exhaustible resource efficiently.[72] Although the authors used some newer mathematics to restate much of what had been previously shown, their main contribution was the proof that uncertainty about future market demand has no effect on the Pareto optimal outcome if suppliers are risk-neutral. They also showed that if suppliers are risk-averse, but society is risk-neutral, then the resource will tend to be underconserved.

In 1976, Stephen Salant published a work that modified the theory of exhaustible resources, by deriving a price-output path of the resource when the industry is characterized by one large cartel, such as OPEC, and several competitive producers.[73] His conclusions were twofold: First, he showed that the above market structure causes a decrease in the exhaustion time relative to a competitive market structure, but an increase relative to monopoly. Second, and possibly more policy-relevant, a disproportionate share of the extra profits due to output restriction goes to the competitors while the cartel ends up being the sole supplier of the resource.

Two comments are worthy of mention: First, by assuming only a certain class of demand curves, his results relating to total exhaustion time are limited. For example, one is left wondering if his "dominant firm" extraction model will always have an extraction time between the monopolistic and competitive outcomes, or if the result is limited to a small class of demand curves. There would be definite policy implications if it could be shown that the "dominant firm" industry was always more socially efficient than a monopolist. The second comment relates to the first. Salant's paper did not address the efficiency attributes of his industrial model, a subject in which economists are presumably interested. It is conceivable that there is no possibility for generalization, but an effort in this direction might have been fruitful.

Notes for Chapter I:

1. W.S. Jevons, The Coal Question, Macmillan, London, 1865.

2. For the moment, the meaning of an "efficient allocation of an exhaustible resource" is not addressed.

3. At the most fundamental level, the pure theory of exhaustion considers that it is profitable, from both a private and a social viewpoint, to limit the current output rate of an exhaustible resource in order that a greater quantity be made available for future use.

4. L.C. Gray, "Rent Under the Assumption of Exhaustibility," Quarterly Journal of Economics, XXVIII (May, 1914), 464-89.

5. H. Hotelling, "The Economics of Exhaustible Resources," Journal of Political Economy, XXXIX (April, 1931), 137-75.

6. O. Herfindahl, "Some Fundamentals of Mineral Economics," Land Economics, XXXI (May, 1955), 131-8.

7. O. Herfindahl, "Depletion and Economic Theory," in Mason Gaffney, (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967, 63-89.

8. W.D. Nordhaus, "The Allocation of Energy Resources," Brookings Papers on Economic Activity, (3: 1973), 529-70.

9. R.L. Gordon, "A Reinterpretation of the Pure Theory of Exhaustion," Journal of Political Economy, LXXV (February, 1967), 274-86.

10. Variational methods refer to methods employing the calculus of variations. See chapter three of this dissertation.

11. F.M. Peterson, "The Theory of Exhaustible Natural Resources: A Classical Variational Approach," unpublished dissertation, Princeton University, 1972.

12. O.S. Goldsmith, "Market Allocation of Exhaustible Resources," Journal of Political Economy, LXXXII (October, 1974), 1035-40.

13. A.C. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay," Journal of Economic Literature, IX (September, 1971), 785-97.

14. R.W. Boadway, "The Welfare Foundations of Cost-Benefit Analysis," Economic Journal, LXXXIV (December, 1974), 926-39.

15. R. Dorfman and N. Dorfman, Economics of the Environment, second edition, Norton, New York, 1977, 402.

16. Ibid., 402.

17. R.M. Solow, "The Economics of Resources or the Resources of Economics," American Economic Review, LXIV (May, 1974), 1-14.

18. User cost is actually a concept drawn from the theory of capital. For a capital asset, marginal user cost is the increment to depreciation caused by an increment of use. See A. Scott, "Notes on User Cost," Economic Journal, LXIII (June, 1953), 368-84.

19. This is not quite the same as seeing market price rise exponentially, Solow tells us. Market price is marginal profits plus marginal costs. Thus market price can conceivably decline while marginal profits rise exponentially, if marginal extraction costs fall through time and if marginal profits are a small part of price. Eventually, however, marginal profits will dominate the value of market price, as price and extraction are connected by the demand curve.

20. H. Hotelling, "The Economics of Exhaustible Resources."

21. The "allocation of resources" can be interpreted at several levels of abstraction. For example, decisions must be made as to the appropriate allocation of resources toward transforming unknown deposits into known ones (exploration), toward the appropriate allocation of mining capital over time, including trained human capital and mineral inventories. Decisions must also be made as to the proper allocation of research and development resources toward the finding of a substitute technology. In the sense referred to in the text at this point, the Pareto efficient allocation of resources refers to the narrow case of allocating a fixed, totally known stock of a depletable resource produced and extracted with a fixed supply of capital.

22. L.C. Gray, "Rent under the Assumption of Exhaustibility."

23. H. Hotelling, "The Economics of Exhaustible Resources."

24. R.L. Gordon, "A Reinterpretation of the Pure Theory of Exhaustion."

25. O. Burt and R. Cummings, "Production and Investment in Natural Resource Industries," American Economic Review, LX (September, 1970), 576-90.

26. H. Hotelling, "The Economics of Exhaustible Resources."

27. F.P. Ramsey, "A Mathematical Theory of Saving," Economic Journal, XXXVIII (December, 1928), 543-59.

28. S.V. Ciriacy-Wantrup, "Taxation and the Conservation of Resources," Quarterly Journal of Economics, LII (February, 1944), 157-91.

29. A. Scott, "Notes on User Cost."

30. D. Carlisle, "The Economics of a Fund Resource with Particular Reference to Mining," American Economic Review, XLIV (September, 1954), 595-616.

31. O. Herfindahl, "Some Fundamentals of Mineral Economics," Land Economics, XXXI (May, 1955), 131-8.

32. Royalties are simply economic profits.

33. Mason Gaffney (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967.

34. O. Herfindahl, "Depletion and Economic Theory."

35. A. Scott, "The Theory of the Mine Under Conditions of Certainty," in Mason Gaffney, (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967, 25-62.

36. W.S. Vickery, "The Economic Criteria for Optimum Rates of Depletion," in Mason Gaffney, (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967, 315-30.

37. S. McDonald, "Percentage Depletion, Expensing of Intangibles, and Petroleum Conservation," in Mason Gaffney, (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967, 269-88.

38. Mason Gaffney, "Editor's Conclusion," in Mason Gaffney, (Ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, Wisconsin, 1967, 333-419.

39. R.L. Gordon, "Conservation and the Theory of Exhaustible Resources," Canadian Journal of Economics, XXXI (August, 1966), 319-26.

40. R.L. Gordon, "A Reinterpretation of the Pure Theory of Exhaustion."

41. It was not clear what Gordon meant by this; for it would appear that firms could maximize current profits because of an uncertain future, yet always need to replenish their mineral source through exploration.

42. F.M. Peterson, "The Theory of Exhaustible Natural Resources: A Classical Variational Approach."

43. R.G. Cummings, "The Economic Theory of Exhaustible Resources: Some Extensions," unpublished dissertation, University of Kansas, 1968.

44. R.G. Cummings, "Some Extensions of the Economic Theory of Exhaustible Resources," Western Economic Journal, VII (September, 1969), 201-10.

45. V.L. Smith, "Economics of Production from Natural Resources, American Economic Review, LVIII (June, 1968), 409-31.

46. R.G. Cummings and O. Burt, "The Economics of Production from Natural Resources: Note," American Economic Review, LIX (December, 1969), 985-90.

47. O.R. Burt and R.G. Cummings, "Production and Investment in Natural Resource Industries," American Economic Review, LX (September, 1970), 576-90.

48. O.R. Burt and R.G. Cummings, "Production and Investment in Natural Resource Industries," 583.

49. F.M. Peterson, "The Theory of Exhaustible Natural Resources: A Classical Variational Approach."

50. N. Vousden, "Basic Theoretical Issues of Resource Depletion," Journal of Economic Theory, VIII (September, 1973), 126-43.

51. Presumably, the conservation motive includes preservation value, option value, existence value, and bequest value.

52. W.D. Nordhaus, "The Allocation of Energy Resources."

53. F.E. Banks, "A Note on Some Theoretical Issues of Resource Depletion," Journal of Economic Theory, IX (October, 1974), 238-43.

54. M. Weinstein and R. Zeckhauser, "Use Patterns for Depletable and Recyclable Resources," Review of Economic Studies Symposium, (summer, 1974), 67-88.

55. R.M. Solow, "Intergenerational Equity and Exhaustible Resources," Review of Economic Studies Symposium, (summer, 1974), 29-45.

56. That is, in effect, neoclassical factors are close enough substitutes for exhaustible ones that society need not depend on exhaustible resources when they run out.

57. P. Dasgupta and G. Heal, "The Optimal Depletion of Exhaustible Resources," Review of Economic Studies Symposium, (summer, 1974), 3-28.

58. R.M. Solow, "The Economics of Resources or the Resources of Economics."

59. N.V. Long, "The Theory of the Mineral Firm and the Optimal Extraction of a Non-Renewable Resource," *Australian Economic Papers*, XIII (December, 1974), 293-7.

60. O.S. Goldsmith, "Market Allocation of Exhaustible Resources."

61. This subject is discussed at length in chapter three, Models III and IV.

62. V. Smith, "An Optimistic Theory of Natural Resources," *Journal of Economic Theory*, IX (December, 1974) 384-96.

63. D.W. Pearce and J. Rose, (Eds.), *Economics of Natural Resource Depletion*, Macmillan, London, 1975.

64. A.J. Surrey and W. Page, "Some Issues in the Current Debate about Energy and Natural Resources," in *Economics of Natural Resource Depletion*, 56-73.

65. G. Heal, "Economic Aspects of Natural Resource Depletion," in *Economics of Natural Resource Depletion*, 118-39.

66. That is, marginal utility is so high in the first period relative to later periods, that the entire use of the resource in the first period would truly maximize social valuation.

67. J.A. Kay and J.A. Mirrlees, "The Desirability of Natural Resource Depletion," *Economics of Natural Resource Depletion*, 145-75.

68. R.M. Solow, "The Economics of Resources or the Resources of Economics."

69. J.A. Kay and J.A. Mirrlees, "The Desirability of Natural Resource Depletion," in *Economics of Natural Resource Depletion*, 171.

70. D.R. Lee and D. Orr, "The Private Discount Rate and Resource Conservation," *Canadian Journal of Economics*, VIII (August, 1975), 351-63.

71. Throughout this dissertation, capital costs are assumed independent of the private interest rate, and instead, are treated as fixed costs.

72. M. Weinstein and R. Zeckhauser, "The Optimal Consumption of Depletable Resources," *Quarterly Journal of Economics*, LXXXIX (August, 1975), 370-92.

73. Stephen W. Salant, "Exhaustible Resources and Industrial Structure: A Nash Cournot Approach to the World Oil Market," *Journal of Political Economy*, LXXXIV (October, 1976) 1079-93.

CHAPTER TWO: A FORMAL PRESENTATION AND EXTENSION OF SOME COMPARATIVE STATICS OF THE PURE THEORY OF EXHAUSTION

A. Purpose of the Chapter

The purpose of this chapter is to provide a conceptualization of the pattern by which the competitive market allocates an exhaustible resource over time. This conceptualization is achieved through the development of three similar models based on a series of simplifying assumptions explained below.

These three models are a formalization and extension of some of the received theory of the mine under certainty. The works which have most heavily influenced the form that these models take are as follows:

In 1914, L. C. Gray[1] developed a verbal presentation of the theory of the mine under certainty. Gray's most important contribution which relates to this chapter is his assertion that it is privately profitable for a competitive mining firm to restrict its output short of that at which price equals marginal cost.

In 1931, Harold Hotelling,[2] using the calculus of variations, proved that a competitive mining industry produced a mineral over time in such a manner that its firms' discounted marginal profits are equal in all periods. Hotelling's proof forms a central part of the three models of this chapter.

Writing in two papers, in 1955 and 1965, Herfindahl[3,4] developed a series of comparative static models of a mining industry, which made the theory of exhaustion accessible to a wide variety of readers. By relying heavily on graphical representation, Herfindahl supported the conclusions of Gray and Hotelling, in addition to exploring the theory of mineral comparative statics. If any one author could be considered the most influential predecessor of this chapter's models, it would be Herfindahl.

Model three of this chapter incorporates the notion, developed by Nordhaus in 1973,[5] of a "backstop technology," a technology of high capital-low exhaustible resource cost, which would be the ultimate substitute commodity for the exhaustibel resource. In this model, the effect on the exhaustible resource sector of a cost decline in the backstop technology is explored.

B. Introduction and Assumptions

Chapter two is the first of two chapters which examine the market allocation pattern of an exhaustible resource over time. The reader will remember that it is important for economists to know how a free market tends to allocate an exhaustible resource because the market is needed as a basis of comparison with a "socially appropriate" allocation. Given such a basis of comparison, a case can be made for or against the desirability of public intervention into the market.

As is the case with most conceptual inquiries into the workings of a market mechanism, certain simplifying assumptions are necessary, and in accordance with the received theory of the mine, the following assumptions will be made:

It is first assumed that the exhaustible resource industry is made up of several firms producing under constant costs, both with respect to output flow and cumulative extraction.[6]

It is also assumed that the industry is characterized by a peculiar form of competition. That is, although individual firms are price takers in the sense that they cannot control price by variation of their own output rates, economic profits do exist. These royalties arise from the scarcity rents accruing to the owners of the scarce mineral. Royalties will be low when the mineral is abundant, but as will be shown, they can be very high when the mineral is nearly exhausted.[7,8]

It is also assumed in this chapter that all the deposits of the mineral are exactly known, fully explored, and individually appropriated.

There is also an assumption made about the existence of forward markets. Specifically it is assumed that there are a set of well-developed forward markets for the mineral. This is assumed because when no forward market exists, even if there is certainty as to the exact size of known deposits, no producer knows the plans of any other producer or consumer. In the absence of this knowledge, producers cannot form accurate expectations of future price patterns. So even though each producer might have a vague idea that the resource price has a long-run upward trend, no producer is able to predict future prices with any certainty. This uncertainty can lead to an unstable market.

For example, assume that a certain set of producers are considering whether to hold their ore or extract it and bring it to market. They feel that price is more likely to rise than to fall. If these producers are averse to risk, they could very well decide to extract now. This decision, of course, drives today's market price down. If this low price persists, not only will the resource be running out faster, but as the resource becomes scarce there will be a general expectation of even higher future prices; producers will react to this by holding supplies off the market, which would send current prices higher yet. Clearly, in the absence of forward markets, the resultant uncertainty could cause long-run market price to oscillate around an equilibrium price, but at any point in time the current price might be as much a reflection of inherent instability as a guide to efficient intertemporal resource allocation.[9]

In this chapter, it is also assumed that there are insurance markets for contingencies. The need for this assumption is best illustrated by an example.

Suppose that there is a 0.5 probability that the breeder reactor will be developed tomorrow, which would produce energy at the cost of one dollar per million BTU; however there is a 0.5 probability that it will not be developed for twenty years. The best available substitute during that twenty year interval has a cost of two dollars per million BTU. Assuming that the market always prefers a lower to a higher price and that either technology could satisfy all demand at its respective price, risk-neutral speculators would establish a market price of $1.50 per million BTU, the expected cost of the substitute for the period.[10] Given this expected price of the resource, its owners would base their production plans on the true $1.50 certainty equivalent for the time period. However, in the absence of such a market for contingencies, if resource owners felt that there was a 0.5 probability of a $1 substitute price and a 0.5 probability of a $2 substitute price, their risk aversion would lead them to act as if the expected price were less than \$1.50.[11] Extraction rates would be biased more to the present than was warranted by the true probabilities. The resource would be underpriced and used up too quickly.[12]

The three models which follow are used to analyze the time profiles of several endogenous variables using mineral demand schedules in a partial equilibrium setting, where demand is assumed constant over time.[13] That is, to the extent that a price change of the resource affects other markets, which, in turn affect the resource market itself, the effect is ignored. In short, the exhaustible mineral market is treated as an isolated unit, independent of the rest of the economy.

Given the above assumptions, it is now possible to outline the process by which unit royalties accruing to owners of an exhaustible resource can be expected to rise at the rate of interest.[14]

First, let us assume that producers initially expect a constant royalty on the marginal unit of output at all points in time. Given that each producer owns a clearly delineated pool of the mineral from which he may draw and that he faces constant costs, then it pays him and others like him to mine all the pool as soon as possible and offer it to the market. Producers will act this way because there is no advantage in postponing production for the future. In fact there is an opportunity cost of postponement equal to the compounded discount rate. Before any unit of mineral is to be kept in the ground, that unit has to yield higher marginal future royalties by a factor of the discount rate. Unless unit royalties are larger by this factor, no producer will allocate any of his output to later periods.

More and more production will be allocated to the earlier periods until the current price begins to fall relative to the expected future price in the forward market.

As the above process continues, the time path of the unit marginal royalties will hover about an exponential growth rate, as producers draw down on their fixed stocks of the resource. As Hotelling stated it, under competition. . . "it is a matter of indifference to the owner of the mine whether he receives for a unit of his product a price (royalty) P_0 now or a price P_0e^{it} after time t."[15] Indeed if the expected time path of per unit royalties rises at a lesser rate than the above rate, perhaps due to too little current supply, there will be a scramble on the part of mineowners to increase their current output rate, rather than earn less than the going rate of return by holding their resource deposits in the ground.[16] This, of course, will drive down present price and royalties relative to those of the future, and tend to return the time path of net profits to an exponential.

It should be pointed out that under the assumptions of this chapter, it is not necessarily the case that each producer continuously decreases his output flow over time. It is just as realistic to think of each as producing at a constant rate, whereupon as the resource becomes "mined out" there are fewer and fewer mines operating, thus putting continuous upward pressure on the mineral price as the supply runs out.[17] Alternatively, one might think of a unit time period of being of the same order of magnitude as the life of a mine.[18]

Thus, it is seen that even under conditions of competitive price-taking and constant costs, both throughout the output range and over time, marginal royalties can be expected to rise at the rate of interest.[19] Hence, it can be said that the phenomenon of exponentially rising marginal royalties is due to both a positive market interest rate and the absolute fixity of the mineral supply.[20]

With this apparatus at hand, let us develop the three models.

C. Model I

In this model, it is assumed that all firms produce a single grade of the mineral[21] at zero marginal cost. That is, the price which faces consumers is made up solely of royalties, all of which accrue to resource owners. It is assumed that no substitute is available at a demand price yielding positive sales. Therefore the exhaustible resource continues in production until market price rises to the point where the quantity demanded is zero, at which time the mineral is totally exhausted, and the industry shifts its factors of production elsewhere.[22]

Keeping in mind the absolute fixity of cumulative extraction, it can be stated:

$$\text{II-1} \quad \int_0^T q(t)dt - m = 0$$

Equation II-1 is used to show that the total rate of production summed (integrated) over all time must equal the quantity available, m. "T" refers to the time at which the mineral is exhausted, and is a variable determined, in part, by total supply. Inserting the equilibrium condition that the supply flow equals the demand flow at all points in time, and that demand price adjusts to quantity put on the market:

II-2 $f(p(t)) - q(t) = 0;$

where demand, $f(p)$ is downward sloping; $f'(p) < 0$.

As mentioned earlier, producers expect per unit discounted royalties to be equal in all time periods for planned production to be forthcoming in all periods. Therefore, in the absence of any external shocks (like new discoveries), per unit royalties must grow at the rate of compound interest. Also, by assuming zero marginal extraction costs, this implies that because the price is composed entirely of royalties, then price increases at the rate of compound interest. That is:

II-3 $R_0e^{rt} - p(t) = 0$, where R_0e^{rt} is the specific form taken by $R(t)$

R_0 = per unit royalties in the initial time period

$R(t)$ = per unit royalties in any given time period, t.

r = discount rate

The final equation shows that at the time of exhaustion, T, price has risen to the point where it has entirely choked off the demand flow, i.e. quantity demanded at time T equals zero.

II-4 $q(T) = 0$

Equations II-1, II-2, and II-3 can be combined into a single equation in implicit form: $F^1(R_0,T,m,r) = 0$, where R_0 and T are endogenous and m,r, are parameters. This combination results in:

II-5 $\int_{0}^{T} f(R_0 e^{rt})dt - m = 0$

Again, II-5 shows that the total quantity of the resource used over time is equal to the given supply, m. However, the expression $f(R_0 e^{rt})$ combines the demand curve, f(p); the price path of II-3, $p = R_0 e^{rt}$; and the exhaustion constraint of II-1.

Note that equation II-4 can also be rewritten as an implicit function of the same variables: $F^2(R_0,T,m,r) = 0$. Using II-2 and II-4, the following expression is obtained:

II-6 $f(R_0 e^{rt}) = 0$

That is, again, price in the last period just forces quantity demanded to zero.

Equations, II-5 and II-6 constitute a system of two equations with two endogenous variables. The endogenous variables are initial per unit royalties,[23] R_0, and time of ultimate exhaustion, T. The two parameters are the discount rate, r, and the total available quantity of the mineral, m;

The Implicit Function Theorem of Mathematics[24] assures us that, <u>under appropriate assumptions of smoothness and the nonvanishing of the Jacobian</u>, II-5 and II-6 can (in principle) be rewritten <u>locally</u> such that R_0 and T appear as two explicit functions, g^1 and g^2 of the parameters, m and r. Thus, from this theorem, it can be said:

II-7 $R_0 = g^1(m,r)$
$T = g^2(m,r)$

More importantly, this theorem shows how to calculate the partial derivatives $\partial R_0/\partial m$, $\partial R_0/\partial r$, $\partial T/\partial m$, $\partial T/\partial r$. According to the theorem, it is valid to perform the following operation, expressed in matrix form:

$$\text{II-8} \quad \begin{bmatrix} \frac{\partial F^1}{\partial R_0} & \frac{\partial F^1}{\partial T} \\ \frac{\partial F^2}{\partial R_0} & \frac{\partial F^2}{\partial T} \end{bmatrix} \begin{bmatrix} \frac{\partial R_0}{\partial m} \\ \frac{\partial T}{\partial m} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \int_0^T f'(R_0e^{rt})e^{rt} & f(R_0e^{rT}) \\ f'(R_0e^{rT})e^{rT} & f'(R_0e^{rT})R_0re^{rT} \end{bmatrix} \begin{bmatrix} \frac{\partial Ro}{\partial m} \\ \frac{\partial T}{\partial m} \end{bmatrix}$$

(Note that the "dt" is not included in the integral expression. This will always be done when the integral is in a matrix, in order to save space.)

The purpose of using the Implicit Function Theorem is to calculate $\partial R_0/\partial m$, $\partial R_0/\partial r$, $\partial T/\partial m$, $\partial T/\partial r$. Equations II-9 and II-10 show the solutions, which may be obtained through use of Cramer's Rule.

$$\text{II-9} \quad \frac{\partial Ro}{\partial m} = \frac{\begin{vmatrix} 1 & f\,(R_0e^{rT}) \\ 0 & f'(R_0e^{rT})R_0re^{rT} \end{vmatrix}}{\begin{vmatrix} \int_0^T f'(R_0e^{rt})d^{rt} & f(R_0e^{rT}) \\ f'(R_0e^{rT})e^{rT} & f'(R_0e^{rT})R_0re^{rT} \end{vmatrix}}$$

$$\text{II-10} \quad \frac{\partial T}{\partial m} = \frac{\begin{vmatrix} \int_0^T f'(R_0e^{rt})e^{rt} & 1 \\ f'(R_0e^{rt})e^{rt} & 0 \end{vmatrix}}{|J|}$$

where, the determinant of "J," in the denominator, is the same term as is in the denominator of II-9.

The above expressions for the partial effect of an increase in the mineral supply on the values for (1) initial period royalties and (2) ultimate exhaustion time are readily calculable, and can be shown to have the following algebraic signs:

$$\text{II-11} \quad \frac{\partial R_0}{\partial m} < 0 \;;\; \frac{\partial T}{\partial m} > 0$$

Equation II-11 agrees with common sense. As there is more of the mineral discovered, initial unit royalties could not be unchanged, for if they were unchanged, then at all points in time, they would be the same as before, due to their increasing at the given rate of interest. Identical per unit royalties at all points in time would imply that at the point when price had risen by enough to choke off all demand, there would still be some of the mineral in the ground.[25] Therefore, per unit royalties shift down at all points in time.

The second half of II-11 also seems to appeal to the intuition. Given that the initial supply was larger, one would expect a larger period of time, T, to pass before the mineral was exhausted. If, in fact, exhaustion time were the same, then the price at that time would not be high enough to exactly discourage all demand. Clearly this could not exist in a competitive market.

Let us now turn to an analysis of the effect on exhaustion time of a change in the rate of discount.[26] Again, the Implicit Function Theorem of mathematics assures us of the validity of a similar matrix expression as II-8. That is, by a similar operation on equations II-5 and II-6 as was done in II-8, it can be shown that:

$$\text{II-12} \quad \frac{\partial R_0}{\partial r} = \frac{\begin{vmatrix} -\int_0^T f'(R_0e^{rt})R_0 e r^{rt} & f(R_0e^{rT}) \\ -f'(R_0e^{rt})R_0 t e^{rT} & f'(R_0e^{rT})R_0 r e^{rT} \end{vmatrix}}{|J|}$$

Given the limitations on the values of the numbers inside the matrix II-12, it can be shown that $\partial R_0/\partial r$ is always negative.

Finally, Cramer's Rule can be used to combine II-5 and II-6 into a form similar to II-8, showing that:

$$\text{II-13} \quad \frac{\partial T}{\partial r} = \frac{\begin{vmatrix} \int_0^T f'(R_0e^{rt})e^{rt} & -\int_0^T f'(R_0e^{rt})R_0 t e^{rt} \\ f'(R_0e^{rT})e^{rT} & -f'(R_0e^{rT})R_0 t e^{rT} \end{vmatrix}}{|J|}$$

With some effort, II-13 can be shown to always take on negative values. Thus, from II-12 and II-13, it can be said that the effect on initial period royalties and the final period of exhaustion of an increase in the discount rate take the following form:

$$\text{II-14} \quad \frac{\partial R_0}{\partial r} < 0 \; ; \quad \frac{\partial T}{\partial r} < 0$$

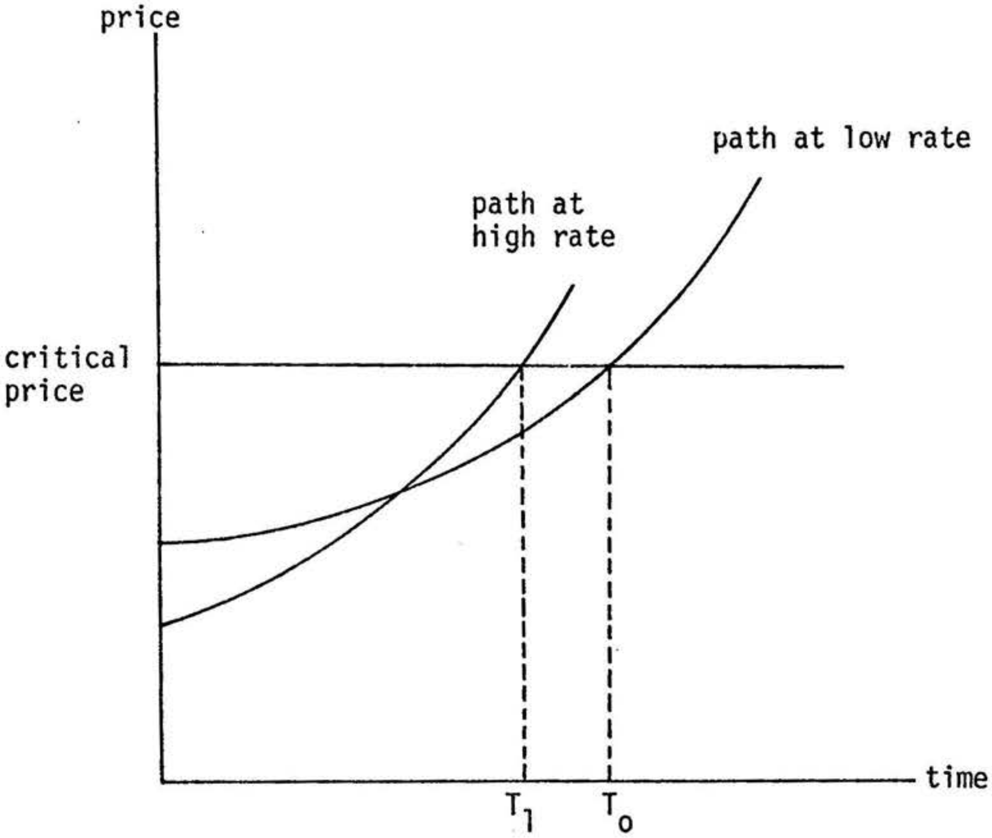
Equation II-14 asserts that with higher rates of discount, the initial value of unit royalties must decrease. This conforms to economic reasoning. As society places a higher value on current consumption, current extraction rates must rise. Due to this higher quantity put on the market in earlier stages, early prices and unit royalties must fall. However prices and royalties could not be lower throughout all time, for then the price in the final period (which must arrive earlier due to the lower price) would be too low to restrict all consumption in that period--a violation of equation II-6, stating that quantity demanded in the final period is zero.

Equation II-14 also says that higher interest rates have the effect of decreasing the period of exploitation. Remember that at some point, the new price path (at the higher discount rate) must cross the old one (at the lower rate). This is shown in figure II-1. It is clear that the period of exhaustion must be earlier because of the higher market price after the intersection point. In other words, if the exhaustion time were the same, then price at that time would be so high that demand would have been choked to zero many time periods before. This is a contradiction.

Summary of Model I

The above analytical presentation has shown the sensitivity of royalties and exhaustion time to changes in the known supply of the mineral and to changes in the discount rate. However the assumption of zero extraction costs and only a single grade of the mineral are very restrictive. The scope for analysis is broadened considerably when these assumptions are dropped in Model II.

FIGURE II-1


price
path at low rate
path at
high rate
critical
price
time
T_1
T_o

D. Model II

Model II is a generalization of Model I. In Model II, it is assumed that there are constant, but positive extraction costs. With this assumption added, the mineral price must be made up of not only unit royalties, but unit extraction costs. Further, it is assumed that there are two grades of the mineral, rather than a single homogeneous grade where all producers face identical costs. In particular, by assuming two grades, it is implied that each grade can be made into the identical end-product but at different costs. Therefore, a unique cost is assigned for each grade of the mineral. In principle, this assumption can be extended to include an arbitrarily large number of grades. Only two are assumed here in order to keep the calculations manageable.

In this model, inquiry is made as to the effect on:

1. Total exhaustion time (the total time for both grades summed together)
2. Total exhaustion time for each grade
3. Initial unit royalties for each grade
4. Initial market price of each grade

of the following parametric shifts:

1. Change in the supply of either grade
2. Change in the marginal cost of producing either grade
3. Change in the discount rate

Before formalizing this model, let us discuss some anticipated results. What are the implications of multiple grades[27] of the mineral? Will the market direct resources toward the simultaneous production of both grades? Given that two identical commodities must command the same price regardless of relative marginal production costs,[28] if both grades are simultaneously produced, then one grade must be yielding its owner a higher unit royalty than does the other grade. However the owner of each grade expects his unit royalties to be growing exponentially. Thus, if both royalties (each of a different value) grow exponentially, then market price must be growing at a different rate for each grade. This, of course, is impossible. Clearly, given the assumptions of this model, simultaneous production of two grades over time cannot occur.[29]

Let us next inquire as to the stable order of exploitation. Figures II-2 and II-3 show two possible orders of exploitation.[30]

Note that at the switch point, $(P-C_b)$ which is equal to R_b, must be less than $(P-C_a)$, which is equal to R_a, because the lower costs of the high grade source necessarily yield higher royalties. Therefore, the slope of the growth path of the low grade royalties, $tR_b e^{rt}$, must be less than the slope of the high-grade royalty growth path, $rR_a e^{rt}$. For this reason, there are only two possible configurations of the order of exploitation, as is shown by Figures II-2 and II-3. In Figure II-2, the producers of the high grade are indifferent to production before or after t_1, because discounted royalties are equal at all points on the price path BB'; however if the low grade producer were to extract first, the price path would begin to follow AA'. Given this state of affairs, the possibility of higher profits would induce high grade producers to dump their output on the market, driving the low-grade producers out of business. The low-grade producers, therefore, can only compete once all the deposits of the high-grade mineral are mined out. The actual order of exploitation induced by the market is shown by Figure II-3.

FIGURE II-2

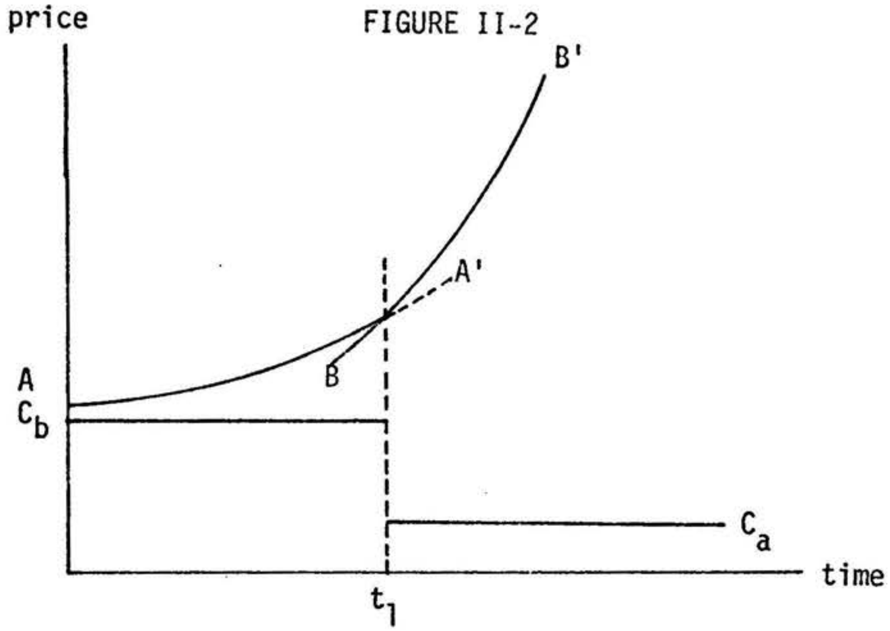
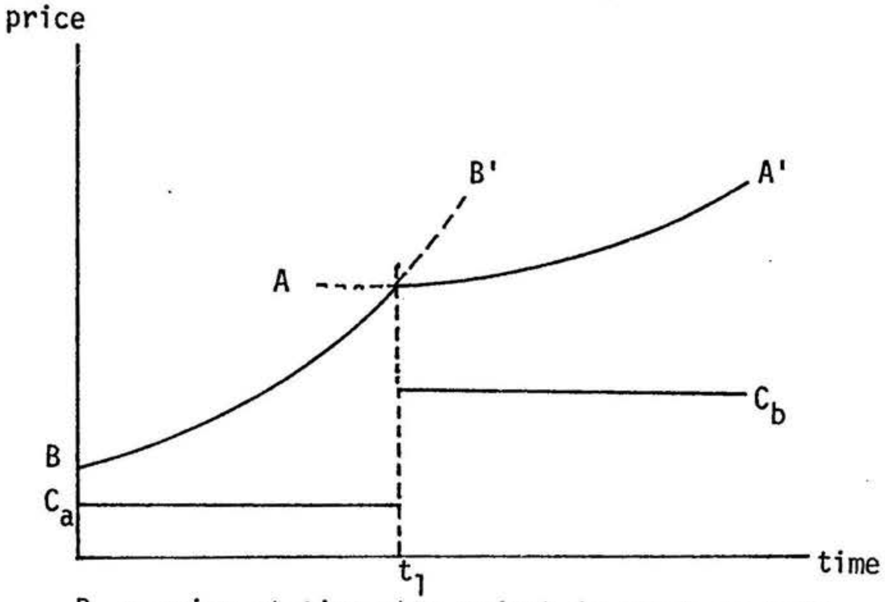


FIGURE II-3



P = price at time when switch in grade is made
C_b = unit cost of low grade
C_a = unit cost of high grade
R_b = unit royalties of low grade
R_a = unit royalties of high grade
t_1 = time at which switch in grade is made

With this background discussion completed, let us begin the formal development of Model II.

Keeping in mind that there are now two assumed grades of the mineral, each with a fixed supply, m_1 and m_2, the total quantity extracted of each grade is equal to those respective supplies:

$$\text{II-15} \quad \int_0^{T_1} q_1(t)dt - m_1 = 0$$

$$\text{II-16} \quad \int_{T_1}^{T} q_2(t)dt - m_2 = 0$$

Note that T_1 and T_2 are the lives of mineral grade one and two respectively. Also note that T_1, T_2, and T, which is equal to T_1 plus T_2, are three endogenous variables; all <u>three</u>[31] can be expected to change given a change in cost <u>or</u> supply of <u>either</u> grade. Equation II-16 states that extraction of the low grade mineral begins at T_1, the exhaustion time of the high grade.

Again, market demand is assumed downward sloping; repeating II-2:

$$\text{II-2} \quad f(p(t)) = q(t) = 0$$

The next two equations are an extension of II-3. They say that the price of each grade includes not only unit royalties of that grade, but unit costs, where unit royalties grow exponentially:

$$\text{II-17} \quad R_{1_o} e^{rt} + c_1 - p(t) = 0$$

$$\text{II-18} \quad R_{2_o} e^{rt} + C_2 - p(t) = 0$$

The R_{i_o} refers to unit royalties at $t = t_o$ for each of the two grades, while the c_i refer to unit costs for the two grades.

The next two equations represent a generalization of II-4: The first equation, II-19, states that the price of the low cost grade equals the price of the high cost grade at the switch point,

$$\text{II-19} \quad R_{1_o} e^{rT_1} + c_1 - (R_{2_o} + c_2) = 0$$

The second equation, II-20, says that the price of the high-cost grade has risen by enough to exactly choke off all demand when both grades are totally exhausted (see equation II-6).

$$\text{II-20} \quad f(R_{2_o} e^{rT} + c_2) = 0$$

where f(.) is the demand function, f(p), and T is the exhaustion time of the high-cost grade.

To calculate the sensitivity of total exhaustion time to exogenous changes, note that:

$$\text{II-21} \quad T - T_1 - T_2 = 0$$

Equation II-21 merely defines T_2, the life of the low-grade source, as being the time at which the industry ends, T, less the time at which the high-grade source is mined out, T_1.

In Model II, prices and royalties are not the same, even though they were in Model I. Thus, any exogenous change will have a separate influence on royalties and mineral price. In general, price equals unit cost plus unit royalty, for either the low- or high-grade sector. This is shown in equations II-22 and II-23.

II-22 $p_{1_o} - R_{1_o} - c_1 = 0$

II-23 $p_{2_o} - R_{2_o} - c_2 = 0$

The last step in laying out Model II is to use the results of equations II-17, II-18, and II-2, to rewrite the more generalized equations II-15 and II-16. The modified equations will be referred to as II-24 and II-25.

II-24 $$\int_{o}^{T_1} f(R_{1_o} e^{rt} + c_1)dt - m_1 = 0$$

II-25 $$\int_{T_1}^{T} f(R_{2_o} e^{rt} + c_2)dt - m_2 = 0$$

Equations II-24 and II-25 provide specific functions of time which can be integrated, rather than the generalized, non-integrable $q_i(t)$ notation.

Equations II-19 through II-25 complete the system of seven equations in seven endogenous variables and five parameters. In implicit form, one can think of the seven equations as taking the generalized form:

II-26 $$F_i(R_{1_o}, R_{2_o}, T, T_1, T_2, P_{1_o}, P_{2_o}; m_1, m_2, c_1, c_2, r) = 0; \quad i = 1, \ldots 7$$

As, in Model I, the Implicit Function Theorem of mathematics is invaluable for the comparative statics analysis required by Model II. In Model II, the theorem enables one to calculate the partial derivative of each of the seven endogenous variables with respect to each of the five parameters. To calculate the thirty-five partial derivatives, the theorem tells us that it is valid to perform the following operation, expressed in matrix form:

$$\text{II-27}\quad \begin{bmatrix} \frac{\partial F^1}{\partial X_1} & \cdots & \frac{\partial F^1}{\partial X_7} \\ \vdots & & \\ \frac{\partial F^7}{\partial X_1} & & \frac{\partial F^7}{\partial X_7} \end{bmatrix} \begin{bmatrix} \frac{\partial X_1}{\partial a_j} \\ \vdots \\ \frac{\partial X_7}{\partial a_j} \end{bmatrix} = \begin{bmatrix} -\frac{\partial F^1}{\partial a_j} \\ \vdots \\ -\frac{\partial F^7}{\partial a_j} \end{bmatrix} \qquad j = 1, \ldots 5$$

where X_i is the ith endogenous variable, and a_j is the jth parameter.

In Model II, as in Model I, Cramer's Rule can be used to solve for the required thirty-five partial derivatives. Leaving the calculations to an appendix,[32] the results are summarized by equations II-28 through II-34.

$$\text{II-28}\quad \begin{bmatrix} \frac{\partial R_{1_0}}{\partial m_1} \\ \frac{\partial R_{1_0}}{\partial m_2} \\ \frac{\partial R_{1_0}}{\partial c_1} \\ \frac{\partial R_{1_0}}{\partial c_2} \\ \frac{\partial R_{1_0}}{\partial r} \end{bmatrix} = \begin{bmatrix} - \\ - \\ - \\ + \\ - \end{bmatrix} \qquad \text{II-29}\quad \begin{bmatrix} \frac{\partial R_{2_0}}{\partial m_1} \\ \frac{\partial R_{2_0}}{\partial m_2} \\ \frac{\partial R_{2_0}}{\partial c_1} \\ \frac{\partial R_{2_0}}{\partial c_2} \\ \frac{\partial R_{2_0}}{\partial r} \end{bmatrix} = \begin{bmatrix} 0 \\ - \\ 0 \\ - \\ - \end{bmatrix} \qquad \text{II-30}\quad \begin{bmatrix} \frac{\partial T_1}{\partial m_1} \\ \frac{\partial T_1}{\partial m_2} \\ \frac{\partial T_1}{\partial c_1} \\ \frac{\partial T_1}{\partial c_2} \\ \frac{\partial T_1}{\partial r} \end{bmatrix} = \begin{bmatrix} + \\ - \\ + \\ + \\ - \end{bmatrix}$$

II-31 $$\begin{bmatrix} \frac{\partial T_2}{\partial m_1} \\ \frac{\partial T_2}{\partial m_2} \\ \frac{\partial T_2}{\partial c_1} \\ \frac{\partial T_2}{\partial c_2} \\ \frac{\partial T_2}{\partial r} \end{bmatrix} = \begin{bmatrix} 0 \\ + \\ 0 \\ + \\ - \end{bmatrix}$$

II-32 $$\begin{bmatrix} \frac{\partial T}{\partial m_1} \\ \frac{\partial T}{\partial m_2} \\ \frac{\partial T}{\partial c_1} \\ \frac{\partial T}{\partial c_2} \\ \frac{\partial T}{\partial r} \end{bmatrix} = \begin{bmatrix} + \\ + \\ + \\ + \\ - \end{bmatrix}$$

II-33 $$\begin{bmatrix} \frac{\partial P_{1_o}}{\partial m_1} \\ \frac{\partial P_{1_o}}{\partial m_2} \\ \frac{\partial P_{1_o}}{\partial c_1} \\ \frac{\partial P_{1_o}}{\partial c_2} \\ \frac{\partial P_{1_o}}{\partial r} \end{bmatrix} = \begin{bmatrix} - \\ - \\ + \\ + \\ - \end{bmatrix}$$

II-34 $$\begin{bmatrix} \frac{\partial P_{2_o}}{\partial m_1} \\ \frac{\partial P_{2_o}}{\partial m_2} \\ \frac{\partial P_{2_o}}{\partial c_1} \\ \frac{\partial P_{2_o}}{\partial c_2} \\ \frac{\partial P_{2_o}}{\partial r} \end{bmatrix} = \begin{bmatrix} 0 \\ - \\ 0 \\ + \\ - \end{bmatrix}$$

where the endogenous variables are defined as follows:

R_{1_o} = initial period royalties accruing to owners of the low-cost (high-grade) source
R_{2_o} = initial period royalties accruing to owners of the high-cost (low-grade) source
T_1 = life of the high grade source
T_2 = life of the low grade source
T = life of the industry
P_{1_o} = initial period price of the high-grade source
P_{2_o} = initial period price of the low-grade source

and where the exogenous parameters are defined as:

m_1 = supply of the high-grade source

m_2 = supply of the low-grade source

c_1 = marginal (equal to average) cost of the high-grade source

c_2 = marginal (equal to average) cost of the low-grade source

r = market rate of interest

Most of the results summarized by the algebraic signs of the partial derivatives in equations II-28 through II-34 would seem to follow one's intuition. However, some of these results have not been discussed in the literature, and would seem to run counter to common sense.

In particular, in equations II-29, II-31, and II-34, note the qualitative effect on initial unit royalties, exhaustion time, and price in the low grade sector--resulting from changes in mineral supply or changes in unit costs in the high grade sector. The effect is shown to be zero. The following would seem to be a reasonable explanation:

Assume, for example, that there is an increase in the expected quantity of high-grade deposits (an increase in m_1). Because of the higher supply, initial price in the high-grade sector must fall. Does this decline in initial price lead to a decreased initial price in the low grade sector? Let us tentatively assume that the initial price in the low-grade sector falls. If this is the case, unit low-grade royalties must be lower at each corresponding point in time, and the time to its exhaustion must be lengthened. This cannot happen, for, since the low-grade supply is unchanged, the mineral would be exhausted before it prices itself out of the market. Contrariwise, if initial price in the low-grade sector increased, it would remain higher at each point in time, and price itself out of the market when there was still something left in the ground. Clearly, initial price, unit royalties, and exhaustion time in the low-grade sector must be unchanged.[33]

Summary of Model II

In Model II, some of the most restrictive assumptions of Model I were dropped. Specifically, Model II allowed for two grades of an exhaustible resource, each with a unique non-zero (constant) marginal cost function, and with a unique life-span of each grade.

Model II developed a series of equations showing how (1) royalties of each grade over time, (2) price of each grade over time, and (3) exhaustion time of each grade is sensitive to changes in (1) the cost of either grade, (2) the supply of either grade, and (3) the market interest rate. The results are summarized in equations II-28 through II-34.

E. Model III

Model III differs from Model II in essentially two ways. First, instead of assuming the existence of two mineral grades, each with a unique marginal cost schedule, it is assumed that there is a single grade of the exhaustible resource which can be substituted by a backstop technology when the resource price rises high enough. In principle, Model III could be extended to several grades, substitutable by a backstop technology. The idea of this "backstop technology" was most fully articulated by William Nordhaus:

> "Ultimately, if and when the transition is completed to an economy based on plentiful nuclear resources (either through breeder or fusion reactors), the economic importance of scarcity of resources will disappear, and capital and factor costs alone will determine prices. This ultimate technology--resting on a very abundant resource base--is the 'backstop technology' and is crucial to the allocation of scarce energy resources."[34]

Second, Model III brings in total discounted net benefits (consumer surplus plus royalties) as a new endogenous variable. The purpose of including net benefits as an endogenous variable is to determine the welfare effect of a change in the cost of the backstop technology.[35]

Thus, Model III inquires as to how:

1. Total extraction time for the exhaustible resource
2. Initial unit royalties accruing to owners of the exhaustible resource
3. Initial price of the exhaustible resource
4. Total discounted net benefits accruing to society attributable to use of the exhaustible resource

are sensitive to a change in the (constant) marginal cost of the backstop technology.

As in Model I, the total quantity of the resource extracted over time must still be equal to its fixed supply. Equation II-5 is repeated for convenience:

$$\text{II-35} \quad \int_{0}^{T} f(R_o e^{rt} + C)dt - m = 0$$

where, as in Model I, f(.) is the demand-price relationship; $R_o e^{rt}$ is unit royalties at time period, t; C is marginal extraction costs of the exhaustible resource, m is its supply, and T is the life of the exhaustible resource.

The next step is new. Note that when the supply of the mineral is fixed and has no substitute, price must eventually rise to the point where the quantity demanded is zero. However, when there is a substitute available at a marginal cost whereby a positive amount will be consumed at each point in time, then it is no longer true that quantity demanded of the exhaustible resource at the period of exhaustion smoothly approaches zero. Instead demand price at the exhaustion time is equal to the cost of the substitute backstop technology.[2,3] This can be written as follows:

II-36 $\quad R_0 e^{rT} + C = C_B = 0$

where $R_0 e^{rT} + C$ equals the market price at time T; and C_B equals the marginal cost of the backstop technology. Note that R_0 and T are endogenous, while C, C_B, and r are parameters.

Let us next repeat II-22 as applied to a single mineral grade:

II-37 $\quad P_0 - R_0 - C = 0$

II-37 says that the exhaustible resource's price equals its unit royalties plus its marginal cost.

Let us finally specify total discounted consumer-plus-producer surplus (net benefits) attributable to use of the exhaustible resource. Following Hotelling,[38] note that the flow of net benefits is an increasing function of the output flow, g(Q), until the level is reached where demand price equals marginal opportunity costs.[39] Output per unit time, of course, is still a function of unit price. With this in mind, one can write:

$$\text{II-38} \quad \int_0^T g(f(R_0e^{rt} + c))e^{-rt}dt - U = 0$$

That is, the integrand of II-35, $f(R_0e^{rt} + c)$, an expression for the output flow at each point in time, Q, can be combined with g(Q), the function for net benefits as a function of output, to obtain a measure of net benefits attributable to the exhaustible resource, over its entire life. This is the meaning of II-38.

Note that dg/df, marginal net benefits of an extra unit of output in period t, equals (P - C), demand price less marginal opportunity cost.[40] U is simply the definitional notation for total discounted net benefits.

Equations II-35 through II-38 complete the system of four equations in four endogenous variables and four parameters. Again, the Implicit Function Theorem of mathematics tells us that each endogenous variable can be thought of as being a function of each of the four parameters.

Using the same techniques as in Models I and II, expressions for the desired partial derivatives of the endogenous variables with respect to the parameters are readily obtainable. The results which are of primary interest are how changes in the cost of the backstop technology affect the endogenous variables, T, exhaustion time; R_0, unit royalties in the first period; P_0 price of the exhaustible mineral in the first period; and U, total discounted net benefits throughout the life of the exhaustible resource. Other sensitivity calculations could be made, but because the backstop technology's cost is the only new variable in Model III, attention is directed solely to the above-mentioned four sensitivities.

As in equation II-26, the Implicit Function Theorem says that the following matrix operation can be performed:

$$\text{II-39} \quad \underset{A}{\begin{bmatrix} \frac{\partial X_1}{\partial C_B} \\ \vdots \\ \frac{\partial X_4}{\partial C_B} \end{bmatrix}} = \underset{B}{\begin{bmatrix} \frac{\partial F^1}{\partial X_1} & \cdots & \frac{\partial F^1}{\partial X_4} \\ \vdots & & \\ \frac{\partial F^4}{\partial X_1} & \cdots & \frac{\partial F^4}{\partial X_4} \end{bmatrix}^{-1}} \underset{C}{\begin{bmatrix} \frac{\partial F_1}{\partial C_B} \\ \vdots \\ \frac{\partial F_4}{\partial C_B} \end{bmatrix}}$$

where the X_1 are the four endogenous variables, T, R_0, P_0, and U. Note that the value of the column vector on the left is that which is being sought. As with Models I and II,[41] the ith element of the desired column vector can be obtained by calculating both the determinant of B and the determinant of B-with-the-ith-column-of-B-replaced-by-the-C-column. The calculations are shown in Appendix C, with the algebraic sign of the vector summarized as follows:

$$\text{II-40} \quad \begin{bmatrix} \frac{\partial T}{\partial C_B} \\ \frac{\partial R_0}{\partial C_B} \\ \frac{\partial P_0}{\partial C_B} \\ \frac{\partial U}{\partial C_B} \end{bmatrix} = \begin{bmatrix} + \\ + \\ + \\ + \end{bmatrix}$$

Note that II-40 says that as the cost of the backstop technology falls, extraction time of the exhaustible resource falls, initial royalties (and the entire time path of royalties) fall, initial price (and its entire time path) falls, and discounted net benefits fall.

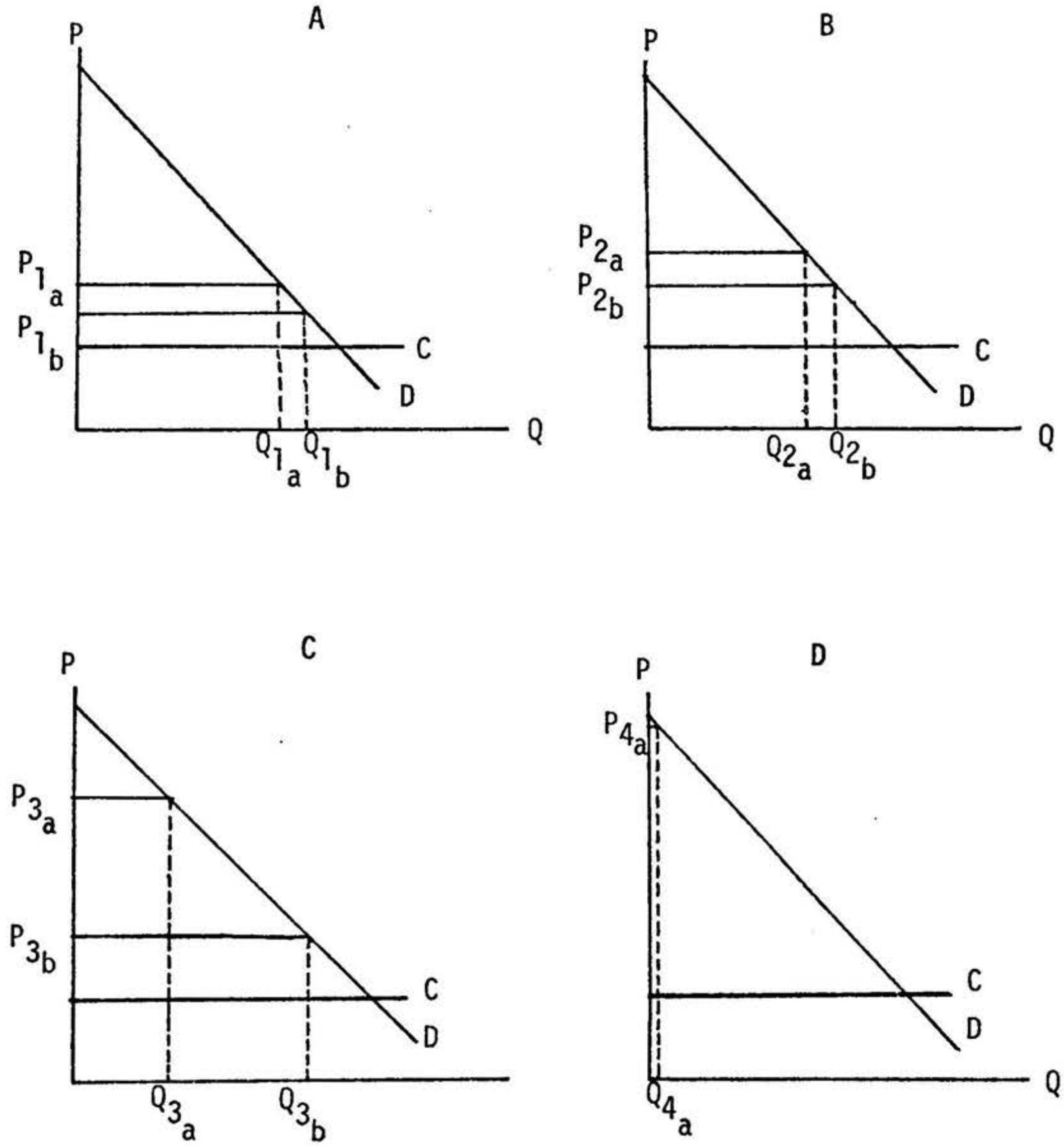
The results of the top three elements in II-40 are worthy in their own right, but the sign of $\partial U/\partial C_B$ warrants some discussion. Note that the algebraic sign of $\partial U/\partial C_B$, the effect on discounted net benefits of a change in the cost of the substitute technology, is positive.

This means that discounted total net benefits (producer-plus-consumer-surplus) fall when the cost of the backstop technology falls. As this may seem unusual, an explanation is in order.

If the current time path of output is such that marginal valuation less marginal opportunity cost[42] grows at less than an exponential rate, then discounted total net benefits can be increased by allocating less output to later periods, and more to present periods. This is because for a given positive discount rate, society is indifferent between marginal net benefits, M, now and Me^{rt} at period t. Therefore, for total discounted net benefits to be a maximum, marginal net benefits must be growing exponentially.[43]

The next question is more difficult to answer. Are all paths where marginal net benefits grow exponentially equally desirable? To answer this question, let us turn to Figure II-4, panels A through D. Assume that the time path of price less marginal cost $(P_{1_a} - c)$. . . $(P_{2_a} - c)$ . . . $(P_{3_a} - c)$ . . . $(P_{4_a} - c)$ is one path of marginal net benefits which grow exponentially. Let us refer to this as "path A." Further assume that $(P_{1_b} - c)$ . . . $(P_{3_b} - c)$ . . .$(P_{3_b} - c)$ is

FIGURE II-4


A
P
P_{1_a}
P_{1_b}
C
D
Q
Q_{1_a} Q_{1_b}
B
P
P_{2_a}
P_{2_b}
C
D
Q
Q_{2_a} Q_{2_b}
C
P
P_{3_a}
P_{3_b}
C
D
Q_{3_a} Q_{3_b}
D
P
P_{4_a}
C
D
Q
Q_{4_a}

another path where marginal net benefits grow exponentially. Let us refer to this as "path B." Assume, that since path A starts at a higher price, it is capable of being sustained for four periods until the mineral is exhausted, while because path B starts at a lower price, production only lasts for three periods. Which of the two exponential paths leads to a higher value of total discounted net benefits?

Focusing attention on panel C, if the marginal unit of output represented by Q_{3b} were shifted out of period three into period four, then period three would lose P_{3b}-less-c net benefits. However, period four would gain P_{4a}-less-c net benefits. The gain, even when discounted by one period, outweighs the loss. This is because P_{4a}-less-c, when discounted by one period, equals P_{3a}-less-c, which is larger than P_{3b}-less-c. Thus, for discounted net benefits to be a global maximum, the quantity of output in the last period must be such that (1) the distance between demand and marginal opportunity cost is highest, and (2) when discounted by one period, yields marginal net benefits equal to those of the previous periods.

The question may arise as to why the market could not start in the first period with marginal net benefits greater than P_{1a}-less-c, then growing exponentially until price rises by enough to choke off demand. The answer is that production could be commanded to follow such a path, but if so, there would still be some of the mineral left in the ground when the price had reached its maximum. This would indicate that the mineral had been mined too slowly for the social good.

Summary of Model III

Thus, it has been shown that, in order for the exhaustible resource's discounted net benefits to be maximized, the price in the final period must entirely choke off the quantity demanded. That is, at any price in the final period less than the choke price, net benefits are not maximized. Further, it can be readily shown that the lower the final period's price, the greater the loss of the exhaustible resource's net benefits. Therefore, since the exhaustible resource's final period's price is equal to the cost of the backstop, decreases in the cost of the backstop lead to a decrease in overall net benefits attributable to the exhaustible resource, and the sign of $\partial U/\partial C_B$ is positive.

The reader must be very careful that he is not led to the false conclusion that a decrease in the cost of the backstop technology yields society negative net benefits. This is incorrect. Model III has only made the statement that total discounted net benefits (consumer-plus-producer-surplus) of the exhaustive resource itself must decline. Naturally, when the backstop's cost falls, consumers face lower prices and receive a higher surplus from that date onward. The concept of declining net benefits accruing to the exhaustive resource is emphasized in order that the net benefits attributable to a cost-reduction of the backstop technology are not overestimated.[44,45]

This completes the series of models which formalize the comparative statics of the mine under certainty. Let us next turn to the inherent efficiency attributes of the afore-mentioned free-market allocation of an exhaustible resource.

Notes for Chapter II:

1. L.C. Gray, "Rent Under the Assumption of Exhaustibility."

2. H. Hotelling, "The Economics of Exhaustible Resources."

3. O. Herfindahl, "Some Fundamentals of Mineral Economics."

4. O. Herfindahl, "Depletion and Economic Theory."

5. W. Nordhaus, "The Allocation of Energy Resources."

6. This assumption is commonly made in the literature, and is not as restrictive as it might first appear. This assumption is consistent with the existence of several different sources of the resource--each extractible at a cost dependent on the characteristics of that source.

7. There are obvious difficulties of imputing some portion of royalties to fixed capital depreciation and some to pure excess profits. Those difficulties are abstracted from in this dissertation.

8. Clearly, it is an oversimplification to express exhaustion as the using up of a clearly-defined quantity of the mineral. Economically, exhaustion can arise due to rising costs, even when there is some of the mineral left in its natural state.

9. In reference to the work of Arrow, Essays in the Theory of Risk Bearing, (Markham, 1971), Houthakker and Solow have argued that if the price of the resource rises more slowly than the rate of interest, then that resource in the ground is a poor investment relative to other forms of capital. By resource owners taking capital losses, this asset nature of the resource will eliminate long-run violent fluctuations in price. See chapter four in this dissertation for more on the subject. Also see Houthakker and Solow, "Comments and Discussion on 'The Allocation of Energy Resources," Brookings Papers on Economic Activity (3:1973), pp. 571-2.

10. Expected cost = .5($1) + .5($2) = $1.50

11. Thus, given a choice between two options, (1) an increment to income, A, with probability, p, and a decrement to income, B, with probability, 1-p, and (2) a certain increment to income, C, a risk-neutral resource owner will be indifferent between (1) and (2) whenever $pA + (1-p)B = C$; the same does not hold for a risk-averse resource owner. Risk averse owners maximize a function of income which increases at a decreasing rate, U(Y), where U'(Y) is positive, and U"(Y) is negative. Thus, the

risk-averse owner, given the same choices, (1) and (2), will find that p(U(A)) + (1-p)(U(B)) is less than U(C); That is, in effect, risk-averse resource owners have "utility" functions which place heavier weights on the probabilities associated with losses than those associated with gains.

12. The specific meaning of an underpriced exhaustible resource will be explained in chapter three.

13. This assumption does some violence to reality, since demand probably shifts out over time due to population and income changes. However, given the knowledge of how demand shifts over time, the effect could easily be incorporated into the models.

14. It is assumed that all deposits are owned privately.

15. Hotelling, "The Economics of Exhaustible Resources," p. 140.

16. Robert Solow, "The Economics of Resources or the Resources of Economics," American Economic Review, Papers and Proceedings, LXIV (May, 1974), 1-14.

17. In the discussion, the term "supply curve" is notably absent. It should be noted that when a resource is exhaustible, the supply curve for the industry (and the firm) is non-existent. This is because output, at each point in time, depends on price at all points in time. Thus, one cannot ask how today's output will change given a change in today's price, for today's price cannot change without future prices changing also. Also, even if that problem were non-existent, a different price path would be traced out by the industry depending on how the demand curve shifted out. Clearly, the notion of industry supply breaks down on both counts.

18. This was suggested by Herfindahl, "Depletion and Economic Theory," p. 64.

19. Of course, market price, which includes per unit costs, will rise at a rate somewhat less than the rate of interest, because per unit cost is assumed constant ove time.

20. In fact, marginal royalties will not rise exponentially if the resource is not in fixed supply. That is, due to the assumed limited supply of the mineral, its rising price cannot bring forth a greater output flow by attracting more firms into the industry. Rather mine owners accrue economic rent, the present value of which equals the discounted value of price less marginal opportunity costs summed (integrated) over all relevant time. This, of course, is not inconsistent with each mine owner being too small to affect the price of the mineral. In fact, we will only see economic rent vanish if the world's supply of the

mineral stock is very large relative to current usage rates. (Herfindahl, p. 75.) Given such a large stock, any short-run economic profits draw new mine owners into the industry until price is bid down to the value of marginal opportunity costs. Then there will be no reason for the mine owner to restrict current output in favor of future output, because the market would not perceive the mineral as being scarce enough to warrant higher price offers for production forthcoming in the relevant future. As a result, with an abundantly available mineral, we would expect each mine owner to operate efficiently as if he were in a typical flow market setting, where no economic rents accrued to the owners.

21. Grade is defined on the basis of costs of production. Thus, by each producer extracting a single grade, it is meant that each firm faces identical constant costs.

22. As this is a model of industry allocation, individual firm behavior is not explicitly considered. Thus, by assuming zero extraction costs for all firms, then "unit royalties" mean those royalties accruing to members of the industry in general.

23. There is no need to know the total time path of per-unit royalties, for once its initial value is known, its value is known for all time, given the discount rate. This is shown by equation II-3.

24. See Appendix A for a statement of the Implicit Function Theorem. Also for a highly readable account of the theory and applications of the theorem, see Alpha Chiang, Fundamental Methods of Mathematical Economics, second edition, (McGraw-Hill, 1974), 198-240.

25. If costs increase with cumulative extraction, the quantity demanded could be zero due to a high price, yet there could be some of the mineral left in the ground. In this simple model, it is assumed that costs do not increase with increases in the total quantity extracted.

26. In the present chapter, there is no welfare meaning attached to the market rate of time preference. Our only concern here is with the effect on the mineral market of a change in the perceived rate of discount. See chapter four for a more complete discussion of the nature and significance of the discount rate from a welfare perspective.

27. "Multiple grades has the following meaning: A given single end-product can be produced at a cost which is dependent solely on the mineral grade."

28. For example, 1000 BTU's worth of coal has equal value regardless of the ore source.

29. To be more exact, ex ante simultaneous production cannot occur. Ex poste production of the two grades could occur simultaneously for a temporary period, but when net royalties are rising exponentially for the low cost producer, they are rising faster than exponentially for the high cost producer. The high-cost producer would realize his error and withhold production from the market until expected unit royalties rose at the appropriate rate.

30. These figures are similar to the ones used by Herfindahl, "Depletion and Economic Theory," p. 73.

31. Note that, although T is not independent of T_1 and T_2, changes in the three will normally each be unique.

32. See Appendix B.

33. It should be noted, though, that even though the initial price and unit royalties are the same in the low-grade sector both before and after the high-grade mineral discovery, exhaustion time of the high-grade source lengthens, as is shown by the positive signs of $\partial T_1/\partial m$ and $\partial T_1/\partial c_1$ in equation II-30. Therefore, <u>discounted</u> initial unit royalties <u>and</u> initial price both decline in the low-grade sector because there is an increase in the waiting period before operations in the low-grade sector are profitable. The subject of total <u>discounted</u> royalties-plus-consumer-surplus-over-the-entire-production period is addressed in Model III.

34. William Nordhaus, "The Allocation of Energy Resources," p. 532.

35. The welfare effect, in this model, does <u>not</u> include net benefits attributable to the backstop technology itself. Rather, it only includes the net benefits attributable to changes in use-patterns of the <u>exhaustible</u> resource.

36. Remember that perfectly functioning futures markets and contingency markets do not allow perverse expectations to develop which might allow some of the ore to remain in the ground.

37. Remember, too, that marginal royalties must rise at the rate of interest, until the resource is mined out, after which marginal royalties in the backstop are zero.

38. H. Hotelling, "The Economics of Exhaustible Resources," p. 143.

39. That is, consumer-plus-producer surplus can be continually increased up to the point where demand price equals marginal opportunity costs.

40. This assumes that there are no market distortions. Market distortions, within the context of applied welfare theory, are discussed in chapter four.

41. This is a restatement of Cramer's Rule.

42. It is assumed that there are no market distortions. Thus, marginal valuation less marginal opportunity cost is equal to price less marginal cost.

43. It should be noted that with constant costs, marginal (or unit) royalties per unit time are the same as marginal net benefits in that time period. Alternatively stated, price-less-marginal-cost equals marginal royalties, which, in turn, are equal to marginal net benefits in that period.

44. Net benefits accruing to users of the exhaustible resource might be thought of as falling when the backstop's cost falls because the resource gets used up at an increasingly "wasteful" rate.

45. In fact the entire issue of consumer-producer surplus benefits, although it may appear to be of only peripheral interest at this point, is brought up again in both chapters three and four.

CHAPTER THREE: WELFARE IMPLICATIONS OF A PERFECTLY COMPETITIVE ALLOCATION OF AN EXHAUSTIBLE RESOURCE

A. Statement of Purpose

The purpose of this chapter is to extend the inquiry of chapter two. Whereas chapter two sought to conceptualize the pattern of competitive extraction, this chapter addresses the higher issue of the desirability of that competitive allocation.

Relative to the stated purpose of this dissertation this chapter is the second of the two chapters which assess the desirability of a market allocation of an exhaustible resource. It is in this chapter where the market allocative response is compared to an appropriately defined efficiency norm.

B. Introduction and Assumptions

As is the case with chapter two, this chapter is not intended to be a self-contained unit. Much of what is discussed in this chapter has been covered in the past literature; this chapter's extensions of the pure theory of exhaustion, although novel, are certainly not path-breaking. Chapter three is included primarily as a conceptual basis for the development of chapter four's welfare loss function. Hence, rather than being an end in itself, chapter three should be considered as a theoretical underpinning for the development of chapter four.

In this chapter, the basic concern is that of comparing the competitive market's allocation of an exhaustible resource with that of an "ideal" or "efficient" allocation. In this respect, this chapter's concept of the "ideal" allocation draws heavily from the notion of Pareto Optimality.

Typically, one thinks of an allocation of resources as being Pareto Optimal if any change in current production or distribution increases one person's net benefits only at the expense of someone else's net benefits. As applied to the analysis of an exhaustible resource, let us say that the allocation of the resource's output is Pareto Optimal if total net benefits (discounted to the present) cannot be increased by any time-reallocation of the exhaustible resource's output.[1] Following Weinstein and Zeckhauser,[2] and others, ". . . the characteristic by which we will measure the optimality of a consumption stream is the discounted sum of consumer-plus-producer surplus."

The methodological approach to this chapter's model construction is similar to that of chapter two, in that the models become progressively more advanced and less restrictive. An outline of the structure of chapter is as follows:

Before formally introducing the models, some space is devoted to an introduction to the mathematics of the calculus of variations, a tool which is used throughout this chapter. Because the calculus of variations is central to the construction of these models, it is covered immediately preceding the models, rather than in an appendix.

The first model is one of a perfectly competitive, price-taking, one-firm industry, which attempts to maximize the discounted value of its total mineral profits over the life of the mine. The model specifies optimizing rules which the firm must follow at each point in the time and also calculates necessary conditions for output to be optimal in the "last" period of production. Here, by a price-taking, one-firm industry, reference is made to a firm which forms an expectation of the likely course of the future price path, assumed independent of its own output decisions; the firm then makes the decision as to its optimal output path based on the above expected price path.[3] Since the single firm <u>is</u> the industry, its output decision is presumed to affect the price path through the market demand relationship. That is, as the firm's output flow begins to fall over time according to its optimal plan, price rises due to the demand relationship, because the mine is the sole member of the industry.[4] Necessary conditions are derived for this output to be optimal at each point during the life of the mine and at the final period of the mine's life.

The second model assumes that the above one-firm industry is controlled by a hypothetical "socialist efficiency planner" whose duty it is to allocate output over time in such a way as to maximize discounted consumer-plus producer surplus.[5] This path of output and price is compared to the path forthcoming from the price-taking one-firm industry which maximizes its profits. It is noted that the profit-maximizing firm depletes the resource too quickly.[6]

Model three generalizes model one by developing the output path forthcoming from an industry composed of several mines, each of which chooses its own optimal output path based on its unique mineral claim and its expectations as to the price path. In model three, it is the summed outputs of all these mines that determines the equilibrium price at a point in time, again through the demand relationship.

Assuming that each firm acts as a profit maximizer, industry price and output paths are then determined.

Model four is a generalization of model two. Again one is asked to envisage a hypothetical efficiency planner which directs each mine's activities in such a way that aggregate discounted producer-plus-consumer surplus is maximized. The planner's only constraint is that he recognize each mine's ownership claim, i.e. that each mine's total physical output over time equals that mine's total mineral claim. Model four concludes that under certain assumptions, the efficiency planner's output plans correspond precisely to that of the industry of profit-maximizing competitors.

This chapter also develops a rather extended discussion of the concept of depletion "user cost," a subject which has been treated lightly in most of the literature on exhaustible resources.[7] The user cost section, appearing between models II and III, develops several geometric illustrations and attempts to show the relationship between typical flow microeconomics and the theory of the mine.

C. A Review of Mathematics: The Calculus of Variations.

If it is desired to choose values of certain variables which maximize a given function subject to a given set of constraints at a point in time, the solution can generally be found by classical programming methods. A standard technique used by economists to solve such problems is the method of LaGrange multipliers. This method is especially desirable because it yields useful information on the sensitivities of the optimal value of the objective function to changes in the constraints. Usually these LaGrange multipliers have important economic interpretations. For example, if one wished to determine the optimal output level of a mineral in a given period, which maximized consumer-plus-producer surplus, the LaGrange multiplier technique could be used.

However, if the problem of concern is the optimal allocation of a mineral over time, where the goal is that of maximizing net benefits across many mutually interdependent time periods, classical programming methods are found to be incapable of rendering a solution.

Suppose one thinks of all possible allocations over time of the exhaustible resource, where each allocation is a specific function of time. Then, associated with each allocation is a specific function of time. Then, corresponding to each of these possible functions of time is a particular number, that number being equal to the time-discounted value of net consumer-plus-producer surplus. Therefore, the rule which associates values of discounted net benefits for any given time path of output can be summarized by the following integral:

$$\text{III-1} \quad U = \int_0^T g(Q(t))e^{-rt}dt$$

T = total time of exploitation

$g(Q) = \int_0^Q (p(q) - c'(q))dq$

$p^o(q)$ = marginal willingness to pay (i.e. market demand) as a function of output

$c'(q)$ = marginal opportunity costs as a function of output in a given time period

e^{-rt} = discount factor applied to net benefits at any point in time, t.

The integral summarized by III-1[8] serves as a rule which associates with each possible function of time, a number, which is the value of the integral. This rule defines a function (called a "functional") in which the elements of the domain are themselves functions of some sort.

The approach which will be used in this dissertation to determine the most desirable time path of mineral output is that of the calculus of variations. Alternative techniques of dynamic optimization include linear programming, where time is divided into discrete units, dynamic programming (the Bellman technique[9]) and the Pontryagin Maximum Principle using optimal control theory.[10] Variational methods are used in this chapter because of their relative simplicity and because many of the results have useful economic interpretations that are analogous to those of classical programming.[11]

If a solution exists to a problem in the calculus of variations, a necessary condition must be satisfied which is analogous to the first order conditions in calculus, that all partial derivatives vanish. This necessary condition is known as the Euler equation of the calculus of variations. The Euler equation is a necessary condition which must be satisfied if the following functional is to have an interior extremum:

III-2 Maximize by appropriate choice of $x(t)$ the following functional "J" $J = \int_{t_2}^{t_1} I(x(t), \dot{x}(t), t)\, dt$

The Euler equation which must be satisfied is as follows:[12]

III-3 $$\frac{\partial I}{\partial x} - \frac{d}{dt}\left(\frac{\partial I}{\partial \dot{x}}\right) = 0$$

where the "dot" indicates the time derivative of the variable, x. There are other necessary conditions which must be satisfied in order that the function, $x(t)$ yield a global maximum for the functional, but it will be apparent from the underlying economic analysis that a maximum and not a minimum is being reached, so these other necessary conditions will not be discussed.

The simplest problem in the calculus of variations assumes that the boundary conditions are given, i.e. that $x(t_0) = x_0$ and $x(t_1) = x_1$. However in the particular problem of the optimal time-allocation of a mineral, both the final time, and the final period mineral output are free to vary.[13] In the calculus of variations, when only a given function, $f(x(t_1), t_1) = 0$ at $t = t_1$, is specified, then, in addition to the solution having to satisfy the Euler equation, one must appeal to a necessary condition which must hold at the endpoints. This necessary condition is called the "transversality condition." The purpose of the transversality condition is to determine the terminal x-value of time, $x(t^*)$, which makes the functional a true maximum. The required transversality condition which determine the optimal endpoints, when those endpoints are free to vary, is as follows:

$$\text{III-4} \quad \left(I - \frac{\partial I}{\partial \dot{x}}\,\dot{x}\right)_{t_1} + \left(\frac{\partial I}{\partial \dot{x}}\right)_{t_1} \left(dx/dt\right)_{f(x(t_1),t_1)} = 0^{14}$$

Equation III-4 says that when, in the final period, the function $x(t)$ is free to take on any point defined by $f(x(t_1), t_1) = 0$, then that point can be found by the rules of that equation.

From the practical standpoint of the theory of the mine, the Euler equation picks out an optimal <u>family</u> of time paths of output, while the transversality condition pins down what output should be in the final period, in order that consumer-plus-producer surplus is maximized. That is, the transversality condition, by picking a single terminal point among all available candidates, picks the unique member of the family of functions defined by the Euler equation.

This completes the necessary review of the mathematics, and should provide a basis for an inquiry into the theory of the mine under certainty.

D. Model I: A Model of a Profit Maximizing Mine's Time Path of Mineral Output

In this model, the optimizing behavior of a profit-maximizing-one-firm mining industry is developed. It is assumed that the mine maximizes the discounted value of its profits subject to the constraint that the total quantity extracted over all periods is limited to its given ownership claim.

For the total quantity extracted throughout time to be fixed, $\int q(t)dt$ must be constant. Just as in classical programming, the calculus of variations can handle constraints on the ojbective function. The method of dealing with this constraint is the one suggested by Intriligator.[15]

In general, given that the functional to be maximized:

$$\text{III-5} \quad J = \int_{t_0}^{t_1} I(x,\dot{x},t)dt$$

is subject to the integral constraint:

III-6 $$\int_{t_0}^{t_1} G(x,\dot{x},t)dt = c$$

where G(.) is a continuously differentiable function and c is a constant, the constraint can be accounted for by introducing the LaGrange multiplier, λ, and defining the functional:

III-7 $$J' = \int_{t_0}^{t_1} (I(x,\dot{x},t) + \lambda(t)G(x,\dot{x},t))dt$$

The Euler equation:

III-8 $$\frac{\partial}{\partial x}(I(.) + \lambda(t)G(.)) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}}(I(.) - \lambda(t)G(.))\right) = 0$$

characterizes the optimal solution. As is the case in classical programming, the LaGrange multiplier is to be interpreted as the marginal value of a slight relaxation of the "K" constraint, dJ/dK, <u>assuming that the optimal time path, $x^*(t)$, is followed, both before and after the relaxation of the constraint</u>.

Assume that the mineowner acts as if he wishes to maximize the functional:

III-9 $$J = \int_{0}^{t_1} (pq(t) - c(q(t)))e^{-rt}dt$$

where p = price; q(t) = output at period t; c(q(t)) = total cost as a function of output at period t; J = total time-discounted net profits.

Further assume that the mineowner's total flow of output over the life of his mine is constrained by his ownership claim to the mineral:

$$\text{III-10} \qquad \int_0^{t_1} q(t)dt = K$$

where K is the total supply of the mineral under his control.

Then according to equation III-8, the mineral constraint can be incorporated into the owner's objective function by introducing the LaGrange multiplier, λ,[16] and defining the functional:

$$\text{III-11} \quad J' = \int_0^{t_1} ((pq(t) - c(q(t)))e^{-rt} - \lambda q(t))\, dt$$

In this case the Euler equation is:

$$\text{III-12} \quad \frac{\partial I}{\partial q} = 0; \ (p = c'(q))e^{-rt} - \lambda = 0$$

Equation III-12 says that the mine should allocate its output over time so that price less marginal cost, i.e. marginal profits, are growing at a rate equal to the marginal user cost[17] times the compound discount rate. Gordon[18] has assured us that this rule holds whether or not price is a function of time. In other words, even if the firm has a price expectation function, any combination of decreasing marginal cost due to a lowering of the output rate, and increasing price, which leads to marginal profits growing at the rate of interest, satisfies the firm's optimizing problem.

The next task is to determine what the firm's optimal output flow should be in its final period of production. The approach used here will be to calculate what the mine's marginal profits should be in the last period, thereby, implicitly yielding an appropriate value of output in that period.[19]

The required transversality condition is as follows:

$$\text{III-13} \qquad (pq(t_1)) - c(q(t_1)))e^{-rt_1} - \quad q(t_1) = 0$$ [20]

Economically, equation III-14 shows that price-less-average cost in the final period should equal marginal depletion user cost. From out previous discussion of the Euler equation, which states that for all periods, price-less-marginal cost is required to equal marginal user cost, it can be said that, in addition to the Euler equation having to be satisfied, the firm must equate average to marginal cost in the final production period. This pair of equations uniquely determines the firm's output path over time.

Of course where marginal cost equals average cost, average profits are at their maximum and average costs are at a minimum. In other words not only does the competitor attempt to equate discounted marginal profits in all periods but he designs his output flow so that when the final production period comes, he is able to mine that remaining quantity at minimum unit cost. Clearly the only way to achieve the largest profits on the final parcel of output is to produce it at minimum unit cost.

It cannot be over-emphasized that both the Euler equation and transversality (end point) condition are necessary to specify a unique time path of mineral depletion. This is because there are an infinite number of time paths which satisfy the requisite Euler equation: any time path of extraction which shows discounted marginal profits equal in all time periods meets that requirement. Naturally if the time at which the mineral is to be totally exhausted is pre-specified, there is only one such path satisfying the Euler equation, and the terminal time condition. However, in this case the exhaustion time is not pre-specified. The firm is free to choose that exhaustion time that globally maximizes its profits.

Referring to Figure III-1, assume that $q_1(t)$, $q_2(t)$, and $q_3(t)$ are three representative time paths all of which extract the same total quantity of the resource. Assume also that each fulfills the constant discounted-marginal-profit-throughout-time requirement. Then it is up to the firm to choose the single $q_i(t)$ which maximizes total discounted profits. Clearly, the terminal production time is a choice variable to the firm. That is, the firm may choose any point on the curve $f(q(t_1),t_1) = 0$ that it wishes.

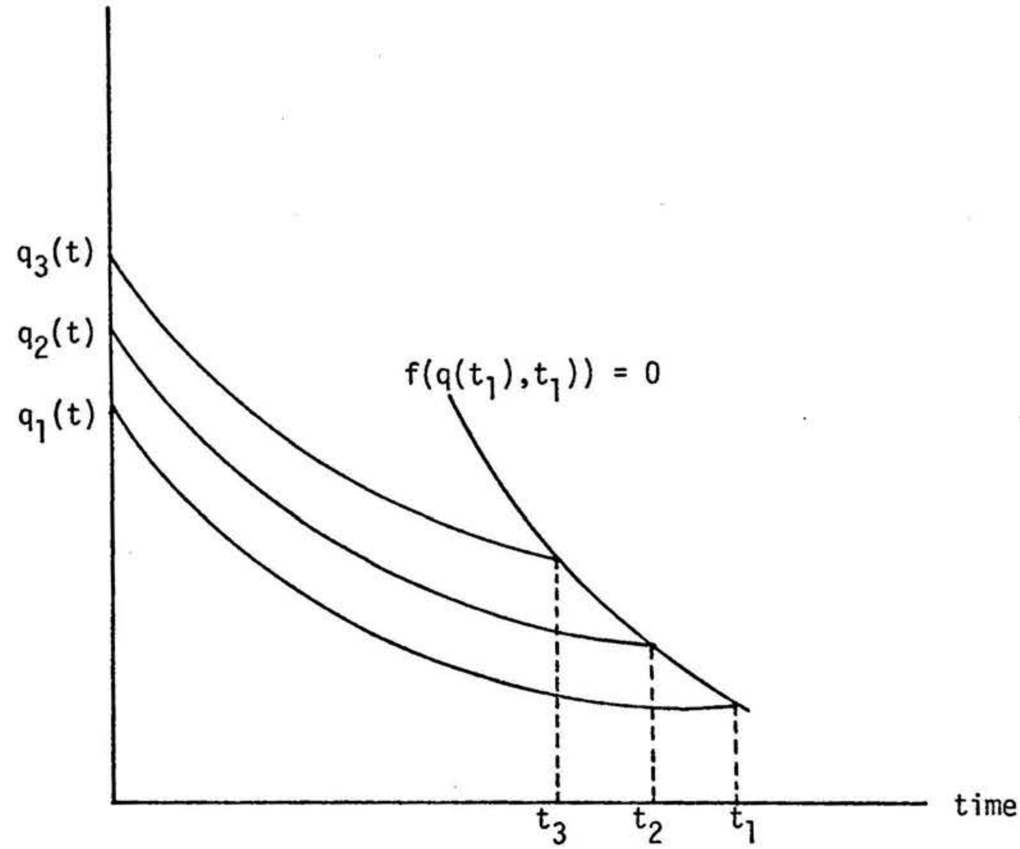
Mathematically, the transversality condition finds the point on $f(q(t_1),t_1) = 0$ which should be selected. In the case where that optimal end-point is given by equation III-14, the profit maximizing firm will choose the single time path of depletion whereby marginal discounted profits are equal in all periods and where unit costs are at a minimum in the final period of production.

An Example:

Up to this point, the output properties of the profit-maximizing-price-taking firm, which operates under a cumulative output constraint have been specified in general terms. Let us now specify a concrete example of the kind of unique output path that is so defined.

According to the Euler equation III-12, it is stated that the individual competitor's output path should be such that marginal profits grow exponentially. This requirement can be expressed algebraically as follows:

FIGURE III-1



$$\text{III-14} \quad MP(t) = \lambda e^{rt} = MP_0 e^{rt}$$

where MP(t) is marginal profits at time, t, MP_0 is marginal profits in the initial period, and λ is a depletion user cost.[21]

However, in general, for the price-taking, one-firm industry, marginal profits can be expressed as a function of output:

$$\text{III-15} \quad MP(t) = p(q(t)) = c'(q(t))$$

Recall that marginal profits are defined as price less marginal cost, where both price and marginal cost are functions of <u>output</u>, itself a function of time.

Keeping in mind that both price and marginal cost are functions solely of output, let us denote the marginal profit function (of output) as h(q). That is:

$$\text{III-16} \quad h(q(t)) = p(q(t)) - c'(q(t)) = MP_0 e^{rt}$$

As will be shown below, it will be necessary to transform equation III-16 into an equation which shows output as an explicit function of time. To achieve this transformation, one must appeal to the inverse function operator, h^{-1}:

$$\text{III-17} \quad q(t) = h^{-1}(MP_0 e^{rt})$$

where h^{-1} is the inverse function of h(.)[22]

With III-17 as an explicit function for q(t), let us show how the Implicit Function Theorem, the Euler equation, and the transversality condition can be used to determine the equilibrium values of MP_0, initial period marginal profits, and T, the exhaustion period, both endogenous variables.

In a manner similar to equation I-1, the total quantity of the mineral used must be equal to the total quantity available. This is shown by equation III-18.

$$\text{III-18} \qquad \int_0^T q(t)dt = \int_0^T h^{-1}(MP_0 e^{rt})dt = K$$

where "K" is the total quantity of the mineral owned by the firm.

Note carefully what equation III-18 has accomplished. This relationship combines <u>both</u> the optimality requirement that marginal profits grow at the rate of interest (the Euler equation) and the exhaustion constraint. Thus, by expressing q(t) as an explicit function of time, equation III-18 incorporates both the Euler equation and the exhaustion constraint. However, equation III-18 is only a single equation in two endogenous variables, MP_0, and T. This bears out what was stated before. There must be another independent equation incorporating the endogenous variables before the firm's output path is unique. That other equation, of course, is the transversality condition. Recall, that economically, the transversality condition requires that, in the mine's final period of output, average cost is a minimum.

If it so happens that average cost always rises, then neoclassical microeconomic theory tells us average cost is minimized when output is zero in the final period. If this is the case, then the transversality condition becomes:

$$\text{III-19(a)} \qquad h^{-1}(MP_0 e^{rT}) = 0$$

where T is the final period of output. However if average cost reaches a minimum at a non zero level of output, then average cost must equal marginal cost in the final period, as specified by the transversality condition:

$$\text{III-19(b)} \quad c'(h^{-1}(MP_0 e^{rT})) = c(h^{-1}(MP_0 e^{rT})) \;/\; h^{-1}(MP_0 e^{rT})$$

where $c'(h^{-1}(.))$ is in functional notation.

Together, equations III-18 and III-19 uniquely determine MP_0, initial marginal profits for the firm, and T, final production period, as functions of the parameters a, the mineral supply, and r, the discount rate. With these two pieces of information, MP_0 and T, the time path of output and price are obtained directly by substituting back into equation III-17 and the demand equation.

Let us constrct a numerical example. Assume that the demand function and cost function are of the linear form:

$$\text{III-20(a)} \quad p(q) = 10 - 5q$$

$$\text{III-20(b)} \quad c(q) = 5q + \tfrac{1}{2}q^2$$

If the firm is to equate discounted marginal profits in all periods, then $p(q) - c'(q)$ must be growing at the rate of interest.

$$\text{III-21(a)} \quad (10 - 5q) - (5 + q) = 5 - 6q = MP_0 e^{rt}$$

If III-21(a) is solved explicitly for q as a function of time, the following is obtained:

$$\text{III-21(b)} \quad q(t) = (5 - MP_0 e^{rt}) \;/\; 6$$

Of course, as emphasized above, III-21(b) only specifies q as a specific function of time when MP_0 has been determined. Before

MP_0 can be known, the mineral supply must be known, as must be the last period's output. For this let us turn respectively to III-18 and III-19(a), to obtain III-22(a) and III-22(b).

$$\text{III-22(a)} \qquad q(t)dt = \int_0^T (5 - MP_0 e^{rt})/6 \; dt = K$$

$$\text{III-22(b)} \qquad (5 - MP_0 e^{rt})/6 = 0$$

After integrating III-22(a) and solving simultaneously with III-22(b), the solution is obtained for MP_0 and T as functions of "K" and "r."

$$\text{III-23(a)} \qquad MP_0 = 5e^{(1-6K/5)r}$$

$$\text{III-23(b)} \qquad T = 6K/5 + 1/r$$

From III-23(a) and III-23(b), it is seen, for example, that a rising mineral supply leads to falling initial-period marginal profits and a longer period of exploitation. A larger discount rate leads to falling first-period marginal profits and a shorter period of exploitation. Both of these results would seem to follow common sense.

From the information in III-23, price and output can be determined at all points in time. This price and output path is summarized by equations III-24(a) and III-24(b).

$$\text{III-24(a)} \qquad q(t) = (5/6)(1 - e^{(1 - r(6K/5 + t))})$$

$$\text{III-24(b)} \qquad p(t) = 10 - (25/6)(1 - e^{(1 - r(6K/5 + t))})$$

Given these results, the model of the one-firm industry is now completely specified.

E. Model II: A Model of the Socially Efficient Output Pattern of a One-Firm Mining Industry

Model II describes the same mining situation as Model I. The difference between the two models is that in Model II, the mine's output decisions are directed by a hypothetical socialist efficiency planner rather than by a profit maximizer. The goal of the socialist efficiency planner is that of maximizing discounted consumer-plus-producer surplus over the life of the mine. The purpose of the model is to compare the hypothetical planner's output choice with that of the free market as a test for the efficiency inherent in the competitive market.

With the aid of Figure III-2, we see that the planner faces a marginal cost schedule, $c'(q)$ and a marginal valuation (demand) schedule, $p(q)$. Subject to the constraint that $\int_0^{t_1} q(t)dt$ is a constant he must choose the output and price at all times (such as t_1) such that the discounted value of the shaded area is maximized when summed over all the periods of the mine's life.

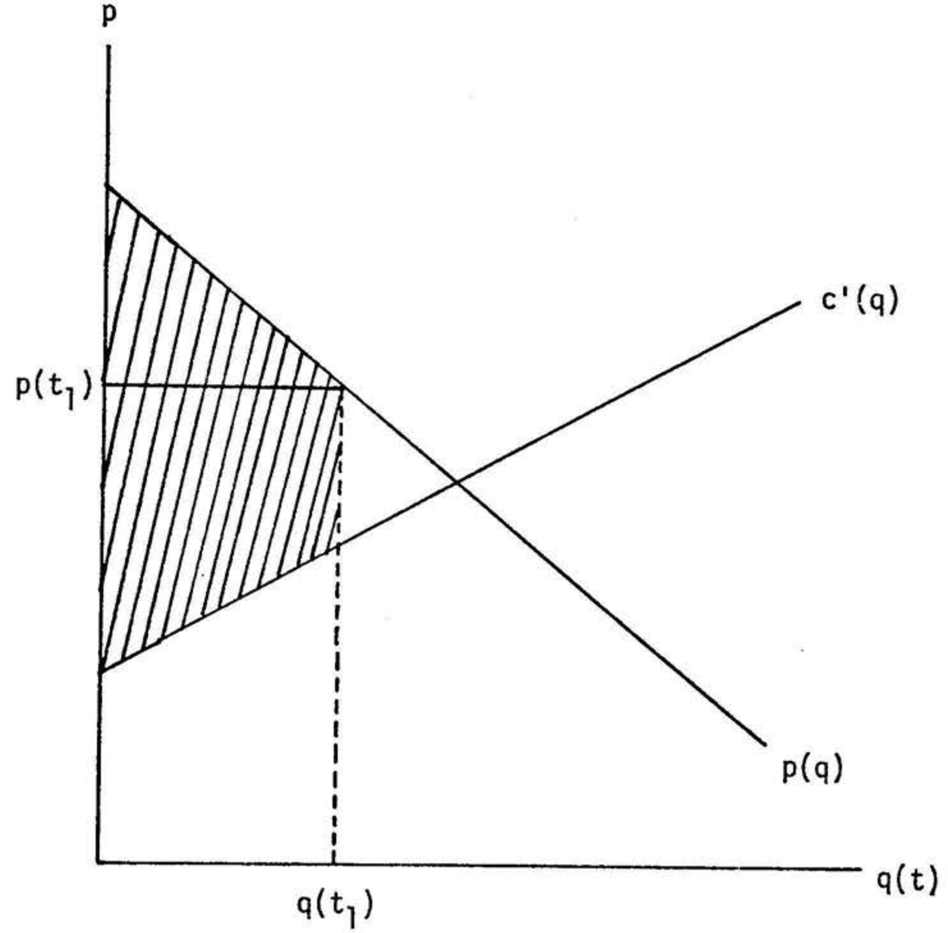
In a given period of time, net benefits, $V(q)$, are given by:

$$\text{III-27} \quad V(q) = \int_0^q (p(s) - c'(s))ds$$

where s is a dummy variable of integration and where $dV/dq = p(q) - c'(q)$; i.e. price less marginal cost at that level of output.

The planner must maximize the total discounted value of $V(q)$ throughout the life of the one-mine industry. Therefore, he wishes to maximize total discounted net benefits, subject to the constraint of a fixed cumulative resource stock. This is shown in equation III-28, where $\lambda\int q(t)$ appears as the mineral constraint.

FIGURE III-2



$$\text{III-28} \quad J' = \int_{0}^{t_1} (V(q(t))e^{-rt} - \lambda(q(t)))dt$$

The Euler equation for the above expression is:

$$\text{III-29} \quad \frac{dV}{dq} e^{-rt} - \lambda = 0$$

where $dV/dq = (p(q) - c'(q))$

The transversality condition is

$$\text{III-30} \quad \frac{V(q(t_1))}{q(t_1))} = \lambda e^{rt_1}$$

where the expression on the left is average benefits in the final time period, t_1.

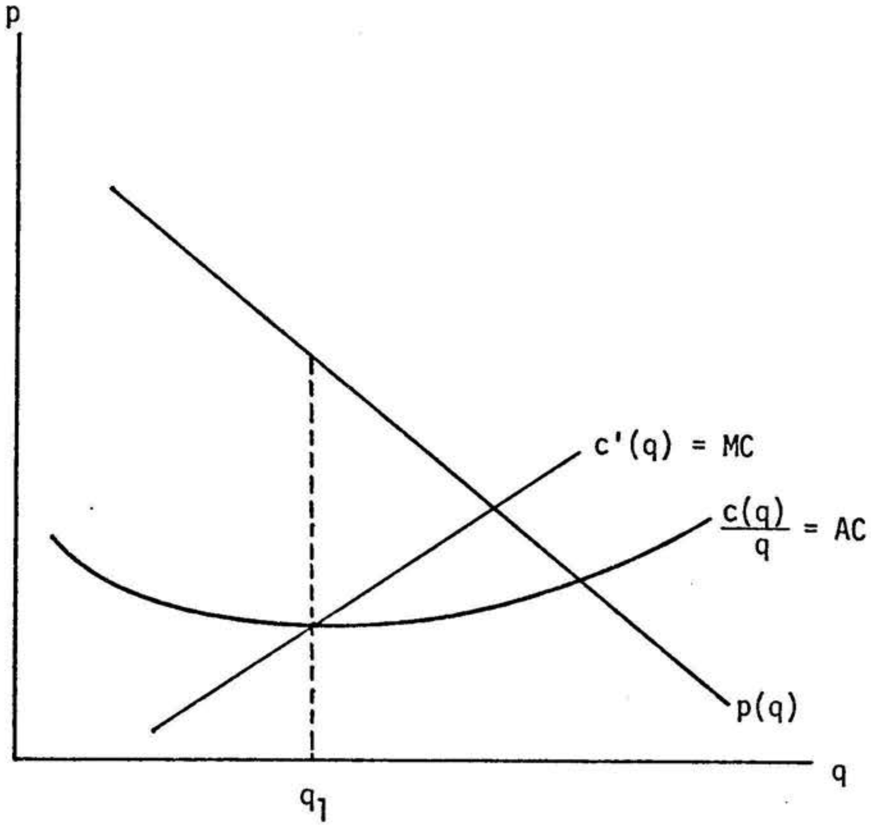
The Euler equation and transversality condition tell us that price-less-marginal cost should be growing exponentially, but that in the final period, average and marginal discounted benefits should be equal, since both are equal to λ in that period.

We now turn out attention to Figure III-3.

Remember that the level of output in the terminal period as determined by the profit maximizer in Model I is at q_1. This is the point where average profits are the highest, since marginal cost equals average cost. The output level chosen by the planner in the terminal period is where marginal benefits equal average benefits, or where unit benefits reach their greatest social value. In Figure III-3, this is shown to occur at a zero level of output.[23]

Therefore, it is seen that the socially optimum terminal[24] level of output is, in general, less than that which would be forthcoming from a profit maximizer.[25] As mentioned in chapter two, this

FIGURE III-3



makes economic sense. In the terminal period, since there is none of the mineral left, unit benefits should be maximized.

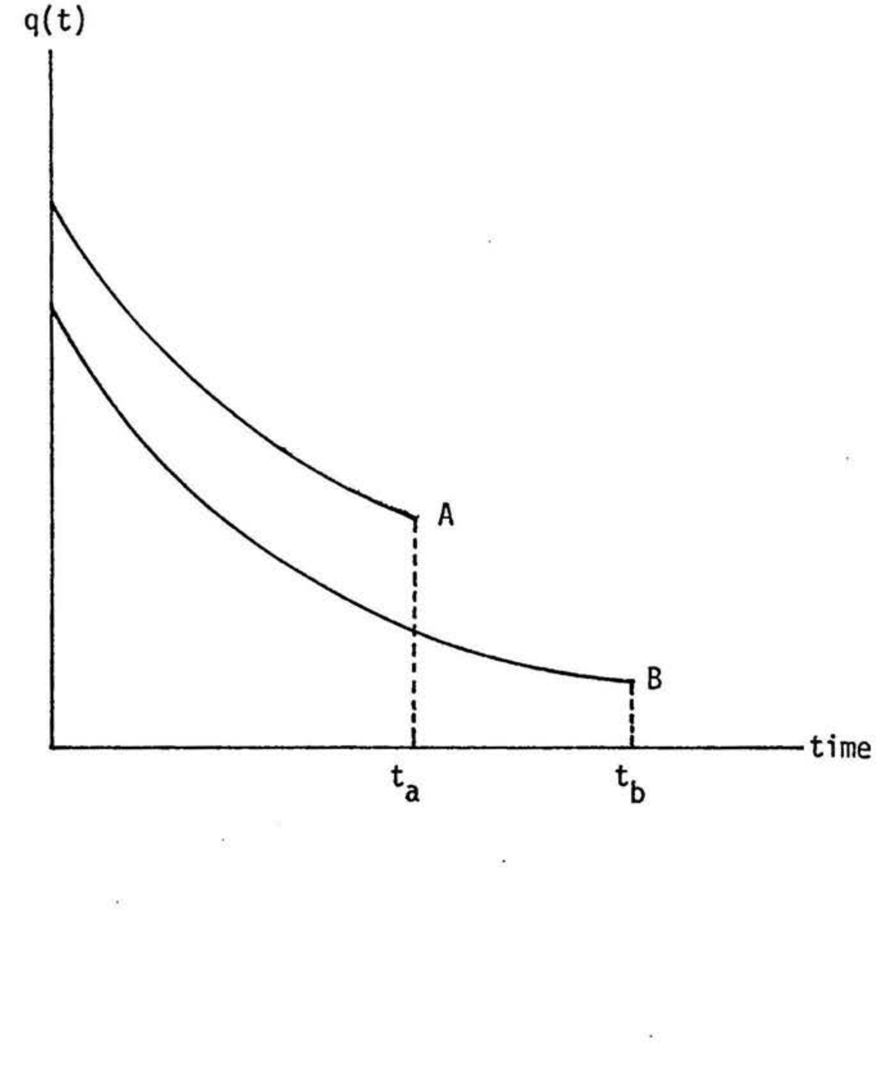
The transversality condition has given us a very important insight. Even though the Euler equation coupled with the market demand curve in Model I gives us the same result as III-29, namely that $(p(q) - c'(q))e^{-rt} = \lambda$; i.e. discounted price-less-marginal cost should be equal in all periods, the two results are not the same, because the λ is different in each case. The two marginal user costs are <u>not</u> the same, and one is told this by the <u>divergent transversality conditions</u>. Since the profit maximizer has a greater terminal output than does the efficiency planner, one can say, in general, that under perfect competition the free market composed of the hypothetical single price-taking firm is inefficient, and that the exhaustible resource is misallocated. Figure III-4 helps illustrate the point.

Time path A represents the path of output under a free-market regime with terminal time equal to t_a. Path B shows how output <u>should</u> be allocated over time in order to maximize discounted consumer-plus-producer surplus. Both paths satisfy the same Euler equation in that discounted price-less-marginal cost is equal in all periods. However path A shows a more rapid depletion of the mineral, and for that reason, is inefficient.[26]

<u>F. A Digression on the Meaning and Significance of User Cost as Applied to Exhaustible Resources</u>

In many economizing problems where an objective function is maximized subject to a side constraint, the technique of LaGrange

FIGURE III-4



multipliers can be used to obtain a solution. Furthermore, once the solution has been obtained, the LaGrance multiplier, λ, is a measure of the greatest possible increase in the objective function attributable to a slight relaxation of the constraint. That is, the multiplier, at the solution, is equal to the marginal value of a slight relaxation of the constraint. The multipliers in Models I and II have similar interpretations.

In Models I and II, λ can be interpreted as the marginal value[27] of a slight relaxation of the constrained, fixed supply of the resource. That is, λ is the marginal contribution to discounted profits attributable to an extra unit of the mineral being discovered. Some authors define marginal user cost as the marginal discounted profits foregone by extracting the unit now instead of at its optimal alternative time in the future.[28] However, upon closer inspection, the two definitions are seen to be identical.

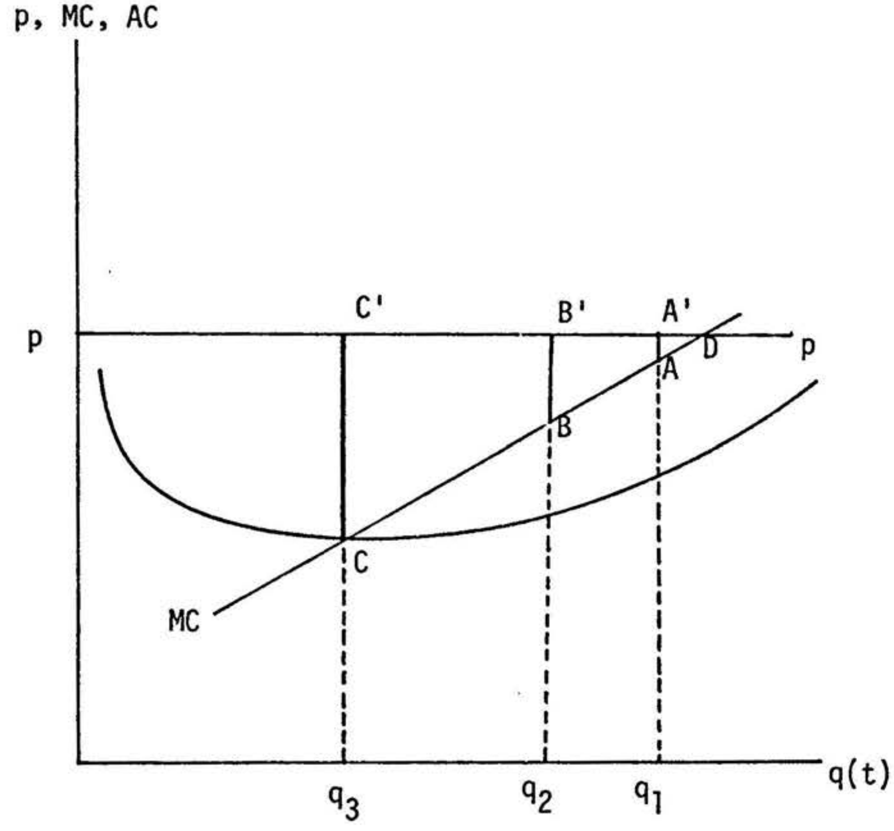
At the margin, if a mine has one million pounds of ore left, the marginal discounted profits given up by extracting the next pound of ore is surely the same as the marginal gain in discounted profits to be had by discovering the million and first pound. Conceptually there is no difference. In fact there is sometimes a third definition given of marginal user cost.[29] This one would define it as the charge in the value of the mine due to operating now rather than picking the best alternative time in the future to mine the unit of output. Clearly, because the value of the mine is the same as the maximum discounted profits accruing to its owner, this definition is also consistent with the other two.

Let us probe somewhat deeper into the user cost concept. First, we must note that once the choice has been made as to the profit-maximizing time path of output, marginal user cost is not a function of time, i.e. $d\lambda/dt = 0$. This is an important assertion to make, for if λ were a function of time, then the rule given by III-12, $(p - c'(q))e^{-rt} = \lambda$ would read "equate marginal discounted profits in each period to marginal user cost in that period." If λ varied by period, then our rule would be of little practical use. The following discussion will show that λ is not a function of time.

If it is true that the optimal time path of output, $q^*(t)$, leads one to produce quantities such that marginal discounted profits are greater in period A than in period B, then total discounted profits could be increased by allocating one more unit of output to A and one less to B. Clearly then, $q^*(t)$ must be sub-optimal and the true optimal time path of output must show marginal discounted profits equal in all periods.[30] That is, λ must be euqal in all periods, and $d\lambda/dt$ equals zero. The optimal time path of output is illustrated below in Figure III-5, where expected price is assumed constant over time.

From the discussion above we see that the firm's production in the assumed three time periods follows the path A-B-C, where output is continuously shrinking in order that marginal undiscounted profits grow exponentially. It should be noted that the value of marginal discounted profits equals AA' in each period. That is, λ must be equal to the distance AA,' or from III-24(a), MP_0, first-period marginal profits.

FIGURE III-5



If time does not influence λ, what does? Does the current output flow affect it? At first sight this would appear to be a reasonable question, but a little thinking will reveal this question to be inappropriate. The current output flow is an endogenous variable to be determined by the mineowner. So to ask if changes in the current output flow affect marginal user cost is really the wrong question, for the current output flow will <u>not</u> be changed, given the parameters of the system. The current output flow will be altered, however, if there are changes in the parameters of the system, so we should be posing the question "how will marginal user cost be affected by parametric shifts?" It is to some of these shifts to which we now turn.

First let us consider changes in the total supply of the mineral. In the extreme case, if the supply of the mineral is very large then the mineowner will hardly be concerned about his current production cutting into the total mineral stock. Fir this reason, output in the earlier periods of production will be carried very nearly to the point where price equals marginal cost. Referring to Figure III-5, the existence of a larger mineral stock will cause earlier periods of production to be carried to points beyond Q_1. Marginal profits in that period, and hence marginal user cost ($\lambda e^{r0} = \lambda$) will be less than AA.'

Let us briefly mention the subject of expectations and their effect on user cost. If future prices are expected to be higher than current prices, the mineowner will be more likely to restrict current output than if the current price is expected to persist indefinitely. Given this expectation, earlier periods of output would be restricted to the left of Q_1 and marginal user cost would rise. Likewise marginal user cost would be higher if production costs are expected to shift down in the future.

Other things being equal, mines with lower production costs will have a higher user cost. This can be seen in two ways. Diagramatically, Figure III-5 shows us that a downward shift in MC or AC will raise AA' and hence increase user cost. Common sense also tells us that when production costs fall, the marginal discounted profits to be realized from a unit discovery of the resource will rise.

Let us next discuss the effect on marginal user cost of changes in the discount rate. At one extreme case, if the discount rate is zero, then today's marginal profits are equally as valuable as tomorrow's; therefore the mine will maximize the total <u>non-discounted</u> profits that could be realized from the resource. This means that production will be carried to the point of minimum average cost each period until the resource is mined out. User cost is increased to a value of cc' (Figure III-5) as each period's output is restricted to Q_3. At the other extreme, if tomorrow's profits are worth nothing (i.e., an infinite discount rate) production will be carried to point D, where marginal profits are zero. Marginal user cost, therefore, declines with a higher discount rate, and the time to exhaustion decreases.

With this discussion on the factors which influence user cost, let us rewrite the optimizing behavioral equation III-12 (The Euler equation).

III-26 $(p - c'(q))e^{-rt} = \lambda(K,c,r)$; $\lambda'(K) < 0$; $\lambda'(c) < 0$; $\lambda'(r) < 0$.

where K equals the total supply owned by the mine; c equals the total cost schedule, a function of output; and r equals the discount rate. $\lambda(K,c,r)$ serves as a reminder to the reader of those parameters which influence user cost.[31]

G. Model III: A Model of the Time Path of Output Forthcoming from an Industry of Profit Maximizing Mines

Model III generalizes the results of a single firm industry. Specifically we assume that there are n firms in the industry, each with a unique cost function, $c_i(q_i)$, each with a unique fixed ownership claim, K_i, and that each will attempt to maximize profits for any arbitrary market price path expected for the future. Of course, the equilibrium market price path depends on the sum of all firms' outputs in each time period and is uniquely determined by the demand schedule in that period. In Model III, the characteristics of the equilibrium price and output paths of the industry will be specified. As in Model I, one might think of equilibrium as being determined by the well-known tatonnement process.

Assume that each of the n firms attempts to maximize its profits subject to its unique ownership claim on the exhaustible resource, and to the equilibrium price path for the industry. Each firm, therefore will act as if to maximize

III-28 $$J = \int_0^{t_i}((p_i(t)q_i(t) - c_i(qi(t)))e^{-rt} - \lambda i q_i(t))dt \qquad i = 1, \ldots n$$ [32]

where t_i is the final period of the ith firm's operation, and the other variables, as before, are $(p_i(t)q_i(t))$, total revenues in period t for the ith firm; $(c_i(q_i(t)))$, total cost in period t for the ith firm; e^{rt}, the discount factor applied by each firm in period t; $_iq_i(t)$, the exhaustion constraint recognized by each firm.

As in the case of a single-firm industry, each firm acts as if to solve its Euler equation, III-29(a), and transversality condition, III-29(b). By doing so, each firm equates marginal discounted profits throughout its unique life, and each produces an output which minimizes unit cost in its final period of output. This is shown below:

$$\text{III-29(a)} \qquad p - c_i'(q_i) - \lambda_i e^{rt} = 0$$

$$\text{III-29(b)} \qquad p - \frac{c_i(q_i)}{q_i} - \lambda_i e^{rt} = 0$$

Note, in equation III-29, that p, equilibrium industry price over time, is determined by total industry output at each point in time. This is shown in equation III-30.

$$\text{III-30} \qquad p - f(q) = 0$$

In equation III-30, q, industry output, is the summed value of the n firms' outputs, $\sum q_i$. This is shown in equation III-31.

$$\text{III-31} \qquad q - \sum q_i = 0$$

In equilibrium, each firm's output plans are based on a price path that is consistent with the summed output plans. Therefore, if the above 2n + 2 equations, III-29 through III-31, are solved, in principle, values can be determined for each firm's output, industry output, each firm's initial period marginal profits, and equilibrium price, as an explicit function of time.[33] That is, in principle, III-29 through III-31 could be solved to obtain the following functions of time:

$$\text{III-32} \qquad \begin{aligned} p &= p(t) \\ q_i &= q_i(t) \\ q &= q(t) \end{aligned}$$

Recalling the previous section's discussion on marginal user cost, it was shown that by solving its transversality condition, a firm equates marginal and average cost in the final period of output. This, coupled with the Euler equation and the resource constraint, uniquely determines the firm's entire time path of output. In fact, the same results hold true for a competitive industry composed of <u>many</u> firms, because each firm acts identically to the single profit-maximizing firm of model I, each <u>perceiving</u> the industry price path as independent of its own output path.

In summary, each firm in the competitive industry, by solving its unique Euler equation and transversality condition, generates output and price paths given by equation III-32.

<u>An Example</u>

Without solving explicitly for an industry of price-output paths, let us add some credibility to the assertions of the previous section. The approach taken is to outline how one might calculate $p'(t)$, the time derivative of the industry's price path; $q'(t)$, the time derivative of the industry's output path, and $q_i'(t)$, the time derivative of each firm's output path. In this outline, it will be shown that, indeed, each firm's marginal profits do grow at the rate of interest, in equilibrium.

Because equations III-29(a), III-30, and III-31 form a system of n+2 equations in n+2 endogenous unknowns, p, q_i, and q, the Implicit Function Theorem tells us that the following matrix operation can be performed:

$$\text{III-32} \quad \begin{bmatrix} dp/dt \\ dq_1/dt \\ dq_2/dt \\ . \\ . \\ . \\ dq_n/dt \\ dq/dt \end{bmatrix} = \begin{bmatrix} 1 & -C_1''(q_1) & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & -C_2''(q_2) & 0 & \dots & 0 & 0 \\ . & & & & & & \\ . & & & & & & \\ . & & & & & & \\ 1 & 0 & 0 & 0 & \dots & -C_n''(q_n) & 0 \\ 0 & -1 & -1 & -0 & \dots & -1 & -1 \\ 1 & 0 & 0 & 0 & \dots & 0 & -f'(q) \end{bmatrix}^{-1} \begin{bmatrix} r\lambda_1 e^{rt} \\ r\lambda_2 e^{rt} \\ . \\ . \\ . \\ r\lambda_n e^{rt} \\ 0 \\ 0 \end{bmatrix}$$

where all variables are as defined in III-29(a), III-30, and III-31.

Next observe that the change in the ith firm's marginal profits over time, can be expressed as

$$\text{III-33} \quad \frac{d(M\pi_i(t))}{dt} = \frac{d(p(t) - c_i'(q_i(t))}{dt} = p'(t) - c_i(q_i(t))$$

where p'(t) is the time derivative of industry price, $c_i''(q_i)$ is the output derivative of the ith firm's marginal cost, and $q_i'(t)$ is the time derivative of the firm's output path.

For illustrative purposes, it will be shown that the marginal profits of firm number one grow at the rate of interest.[34] Given such a two-firm industry, let us use Cramer's Rule to solve equation III-32 (reduced to a system of the four engogenous variables, $q(t)$, $q_1(t)$, $q_2(t)$, and $p(t)$). This is shown in equation III-34.

$$\text{III-34} \qquad dp/dt = \frac{f'(q)(r\lambda_1 e^{rt}\, c_2''(q_2) - r\lambda_2 e^{rt}\, c_1''(q_1))}{-(c_1''(q_1)\, c''(q_2) + f'(q)(c_2''(q_2) + c_1''(q_1))))}$$

$$dq_1/dt = \frac{(r\lambda_1 e^{rt}\, c_2''(q_2) + f'(q)(r\lambda_2 e^{rt} - r\lambda_1 e^{rt}))}{-(c_1''(q_1) + c_2''(q_2) + f'(q)\, c_2''(q_2) + c_1''(q_1)))}$$

After several steps of manipulation, it can be shown that III-34 can be rewritten as:

$$\text{III-35} \qquad p'(t) - c_1''(q_1)q_1'(t) = r\lambda_1 e^{rt}$$

where λ_1, marginal profits in the first firm's initial period, is determined by that firm's optimal choice of its terminal-period output. (See model I's example.)

Equation III-35 shows the desired result, namely that along the industry's equilibrium price path, the growth rate of firm one's marginal profits, $d(M\pi_1(t))/dt$, is exactly equal to the growth rate of $\lambda_1 e^{rt}$.

<u>Summary of Model III</u>

Model III extends the results of Model I. Model III assumes the existence of a competitive industry, made up of <u>n</u> firms, each having claim to a unique, well-defined deposit of an exhaustible resource. Assuming that <u>each</u> firm acts as if to maximize the time-discounted value of total profits attributable to its stock of the exhaustible resource, then (1) each firm chooses an output path that equates marginal discounted profits in all periods and minimizes average cost in the final period, and (2) the results of all firms' actions give rise to an equilibrium industry price path, along which each firm maximizes its profits.

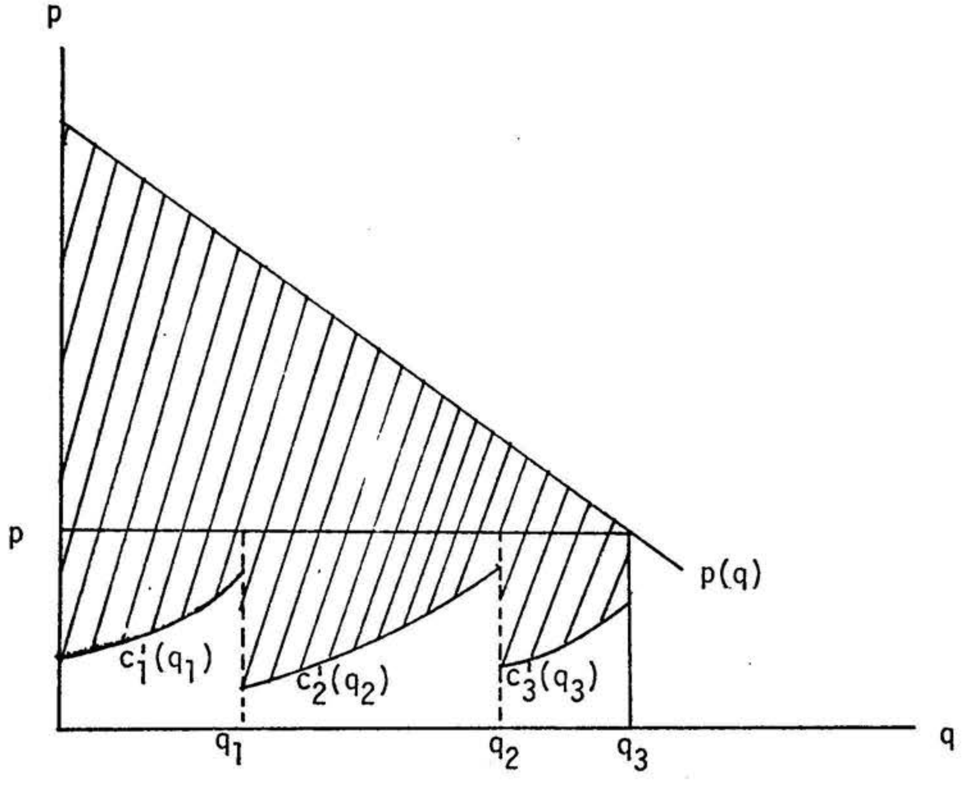
H. Model IV: A Model of a Socially Efficient Mining Industry

Model IV is a direct comparison of the results of Model III. Model IV describes the output and price paths of an industry of n firms, whose output decisions are determined by an efficiency planner. In this model, it is shown that, given some simplifying assumptions, there will be a tendency to deplete the exhaustive resource at the efficient rate in a competitive mining industry.

It should be emphasized at the outset that this result is not new to the literature of the economic theory of exhaustible resources. It has been stated in the past--by Hotelling,[35] Peterson,[36] Goldsmith,[37] and probably others. The difficulty with these authors' proofs was that they typically ignored the significance of the transversality condition.[38] As was emphasized in models I, II, and III, _both_ the Euler equation _and_ transversality condition must be the same under both the market and efficiency regimes in order that both yield equal outcomes. Model IV does just that. It shows that the competitive industry and the efficiency planner face the same Euler equations and, properly interpreted, the same transversality conditions for each of the n firms.

Total surplus, in a given time period is given by the area under the demand curve less the summed areas beneath all firms' marginal cost curves, (shown in Figure III-6). It is this total surplus summed over the life of all firms that the efficiency planner wishes

FIGURE III-6



to maximize.[39] Mathematically, the planner perceives the following as a measure of net benefits at a point in time, t:

$$\text{III-36} \qquad V(q_1, \ldots q_n) = cs(q) - \sum c_i(qi)$$

where $\sum q_i = q$; cs(q) equals gross consumer surplus, equal to the total area beneath the demand curve, up to the level of output flow, q, the derivative of which is p(q), the demand price; $c_i(q_i)$ = total cost of the ith firm, as a function of its output, q_i. Next, following Goldsmith,[40] it can be said that the efficiency planner wishes to allocate output of the mines over time in such a manner as to maximize:

$$\text{III-37} \quad J = \int_0^T v(q_1, \ldots qn)e^{-rt} - \int_0^{t_1} \lambda_1 q_1 - \ldots - \int_0^{t_n} \lambda_n q_n$$

where λ_i is the "social" user cost of an increment to firm i's resource stock, i.e. the maximum marginal contribution to <u>net benefits</u> of an extra unit of the resource stock discovered at the ith mine site; t_i is the final period of operation of the ith mine; and T is the final period of operation of the industry itself, equal to the greatest of the t_i.

By maximizing III-37, the efficiency planner is faced with the following <u>n</u> Euler equations:

$$\text{III-38} \quad \frac{\partial v}{\partial q_i} = \lambda_i e^{rt} \qquad i = 1, \ldots n$$

where $\partial v/\partial q_i$, the increment in total welfare of a unit increase in the ith mine's output flow at a point in time, equals price less marginal cost, $p - c_i'(q_i)$.

The efficiency planner is also faced with the following n transversality conditions which require that, for each mine, $J(t_i) = 0$, where J is the expression in III-37, and t_i is the final period of the ith mine's operation:

$$\text{III-39} \qquad cs(q_i(t_i)) - c_i(q_i(t_i)) - \lambda_i e^{rt_i}$$

In equation III-39, it would appear that given a very large number of mines, $cs(q_i(t_i))$, the contribution of the ith mine to consumer surplus in its final period of output, should be interpreted as price times the output of that mine in that period.[41]

Given such an interpretation, III-45 says that, in effect, $(p-c_i(q_i(t_i)))/q_i(t_i)$ should be equal to $\lambda_i e^{rt_i}$. That is, each mine's final production period should have average cost equal to marginal cost, a result identical to that of Model III of the profit-maximizing series of competitive firms.[42]

A Comparison of Results

A note in the comparison of Models I-II as opposed to Models III-IV is worthy of mention, since the results would seem to conflict. Recall, that in Model II, it was shown that a single mine operated by an efficiency planner exhausts a mineral more slowly than the same mine subject to the control of a price-taking profit maximizer. The reason for the difference was shown to lie in the divergent end-point conditions: The competitor maximizes average profits in the final period, while during that terminal period, the efficiency planner maximizes average benefits, part of which include the area under the demand curve that the competitor cannot capture. Unless average cost rises throughout its range, the competitive solution is one of premature resource exhaustion.

Contrast this result with that of the competitive industry, which appears to exhaust the resource efficiently over time. In a competitive industry made up of many small firms, no single firm's output can be considered as contributing to consumer surplus, over and above its contribution to market price. This is because $\partial p/\partial q_i$, the marginal effect of the ith firm's output on industry price, is virtually zero, even if all firms are under the control of a single perfectly discriminating monopolist or efficiency planner. In the hypothetical one-firm industry, however, the requisite end-point efficiency condition of average benefits being equal to marginal benefits leads to a different output than the competitive market outcome, where average equals marginal profits. However, there does not appear to be any such divergence between the two regimes for the many-firm industry. This is because in a many-firm industry, the output of a single firm has a contribution to average benefits equal to that firm's contribution to its owner's average profits; namely both are equal to

III-40 Average benefits (t_i) = average profits (t_i) = $p - c(q_i)/q_i$[43]

Notes for Chapter III:

1. This definition of Pareto Optimality assumes that a discounted dollar's worth of resource output accruing to a member of a future generation is equal in value to a dollar's worth accruing to a member of today's generation. This is discussed in chapter four.

2. M. Weinstein and R. Zeckhauser, "Optimal Consumption of Depletion Resources," Quarterly Journal of Economics, LXXIX (August, 1975), 373.

3. This, of course, is a naive model because the single firm in question would undoubtedly quickly realize that its output influences market price, and as a result it would behave as a monopolist. However the model is a good starting point as it leads to a similar model of industry behavior.

4. To make this acceptable, one can visualize the adjustment process taking place through a tatonnement mechanism. The auctioneer would call out price time paths and get back output responses from the firm. The auctioneer would order production to commence when the time paths of demand and supply coincided. That is, the equilibrium solution is reached when the mine's output over time gives rise to a price path which corresponds exactly to its expected price path.

5. The notion of a socialist efficiency planner choosing an ideal allocation of resources, the outcome of which is compared to a free market allocation, was probably generalized most elegantly in F.M. Bator," The Anatomy of Market Failure," Quarterly Journal of Economics, LXXII, (August, 1958), 351-79.

6. It could be said that this result is because of the naive price-taking assumption of the model; however if the same industry were controlled by a monopolist, it has been shown in the literature that the resource is still depleted suboptimally.

7. For a notable exception, see G. Canarella, "Optimal Policies for Depletable Resources," (unpublished dissertation, University of Virginia, 1973).

8. This is the equation introduced by Hotelling, "The Economics of Exhaustible Resources," 143.

9. R.E. Bellman, Dynamic Programming, Princeton University Press, (Princeton, N.J.), 1957.

10. C.S. Pontryagin, et al., The Mathematical Theory of Optimal Processes, Interscience Publishers, New York, 1962.

11. For an excellent discussion of the relationship between classical programming and the calculus of variations, see M. Intriligator, Mathematical Optimization and Economic Theory, (Prentice-Hall, 1971), especially chapters three and twelve.

12. The Euler equation is merely stated here. For a highly readable derivation, see M. Intriligator, 308-310.

13. In principle the initial period is also free to vary, as noted in Gordon, "A Reinterpretation of the Pure Theory of Exhaustion." However, it is assumed in this chapter that the time when production commences is $t = 0$.

14. In many cases, the curve on which x can lie in the final period $f(x(t_1),t_1),t_1) = 0$, is given as a part of the problem. However, this is not always the case. For example, if it is desired to determine the distance-minimizing function, which starts at the origin and ends on any point, such that the area beneath the function is constant, there is no explicit terminal curve given. However, upon solving the Euler equation for the constrained minimization problem, the terminal curve is determined to be a rectangular hyperbola. According to the transversality condition, the required straight line function should pass through the hyperbola perpendicularly.

15. M. Intrilligator, Mathematical Optimization and Economic Theory, 318.

16. It will be shown that in this special problem under consideration, λ is independent of time.

17. Marginal user cost is the marginal profits attributable to an incremental relaxation of the mineral supply constraint and is equal to λ in equation III-12. This subject is covered in greater depth after model II.

18. R. Gordon, "A Reinterpretation of the Pure Theory of Exhaustion," 279.

19. Hotelling, "The Economics of Exhaustible Resources," discussed the significance of the transversality condition, but rather than applying it to the competitive case, he only explored its implications for a monopolist.

20. Normally, it would have been necessary to find the function $f(q(t_1),t_1) = 0$, before the transversality condition could be applied. However, since $\dot{q}$, the time-derivative of q, does not appear in III-11, then from III-4, it is necessary only that the function in the integral of III-11 equal zero in the last period. Hence, the logic of III-14.

21. Alternatively, λ is the maximum marginal contribution to the objective function attributable to a slight relaxation of the mineral stock constraint.

22. The function h(.) must be monotonic for III-18 to be valid.

23. Even if marginal cost declines at early levels of output, marginal-benefits-equals-average-benefits <u>must</u> occur at a lesser level of output that q_1; i.e. one is comparing the areas $\int p(q) - c'(q)/q$ to $(p(q) - c'(q))$, or $\int p(q) - c'(q)$ to $(q(p(q) - c'(q))$. At q_1, $q(p(q) - c'(q))$ equals total profits. However, $\int p(q) - c'(q)$ includes total profits <u>plus</u> consumer surplus. These two cannot be equal unless output is less than q_1, the profit maximizer's terminal level of output.

24. The socially optimum time path of output maximizes discounted consumer-plus-producer surplus.

25. There is an exception to this rule. If average cost is always rising, as was shown in Model I, the competitor will produce a zero level of output in the final period. In such a case, the competitor achieves the socially optimal allocation of the resource over time.

26. This "inefficient result" is probably due to the unrealistic assumptions of the one-firm price taking industry, and as such should not be taken too seriously. However, the economic importance of the transversality condition should not be overlooked, as it plays an important role in the industry models discussed in Models III and IV.

27. In Model I, value is with reference to total private discounted profits. In Model II, it is with reference to total discounted net benefits. Thus, in this section, "value" refers to the objective function of either Model I or Model II, depending on the context.

28. See Anthony Scott, "Notes on User Cost," <u>Economic Journal</u>, LXIII, (June, 1953), 368-84.

29. For an excellent discussion of several types of user costs, especially as they apply to the petroleum industry, see Paul Davidson, "Public Policy Problems of the Domestic Crude Oil Industry, <u>American Economic Review</u>, LII (March, 1963), 85-108. Keynes was probably one of the first writers to discuss user cost extensively; however his concern with the concept lay in the theory of capital. See <u>The General Theory of Employment, Interest, and Money</u> (New York: Harcourt, 1936), 66-73.

30. This logical result has been in the literature for several years. It appears to have been first formulated by L.C. Gray, "Rent under the Assumption of Exhaustibility," especially pp. 475-477.

31. Recall, that in equilibrium, the maximum discounted increment to total discounted profits attributable to an incremental "find" of the mineral, is equal to first-period marginal profits. Thus, in equilibrium, λ , marginal user cost, is equal to first-period marginal profits; this can be seen by recalling the numerical example of Model I, where, in equilibrium, first-period marginal profits, MP_o, were shown to depend on the mineral supply, K; the interest rate, r; and the marginal cost schedule.

32. The subscript, i, associated with price indicates that each firm may have a unique perception of the time path of price. However this subscript is removed in equations III-29(a) and III-29(b), because the equilibrium price is assumed to be determined by the tatonnement mechanism, corresponding in reality, to a set of perfectly functioning futures and contingency markets.

33. Only 2n+1 of the equations are independent, since the sum of the q_i's must be equal to q, industry output.

34. This can be readily generalized to the marginal profit function for the ith firm in an n-firm industry.

35. H. Hotelling, "The Economics of Exhaustible Resources," 143.

36. F. Peterson, "The Theory of Exhaustible Natural Resources: A Classical Variational Approach."

37. O. Goldsmith, "Market Allocation of Exhaustible Resources."

38. Peterson derived results which showed that the transversality condition for a competitive industry was, in general, different from that of an efficiency planner. However, Peterson did not comment on this discrepancy. See Peterson, p. III-13.

39. One might think of the efficiency planner as wishing to maximize the same figure as does a perfectly discriminating monopolist. However, the area between the demand curve, and the price line is the monopolist's profits, rather than consumer surplus.

40. O. Goldsmith, "Market Allocation of Exhaustive Resources," 1036.

41. This would be as opposed to including some of the area beneath the demand curve and above market price as a part of the contribution of that mine. That is, in a single mine's final period of output, only a negligible amount of consumer surplus is being generated; hence that mine's average contribution to consumer surplus is $pq_i(t_i)$, because $cs(q_i(t_i))$ has no economic meaning other than pq_i.

42. This conclusion assumes that $cs(q_i(t_i))$ is really equal to $pq_i(t_i))$, generally a safe assumption when a given firm's output is small relative to the industry. However, as the industry's life comes to an end, there are progressively fewer firms left, and that assumption is violated. Perhaps, in these later periods, the efficiency planner would restrict output more than would the industry of competitive profit-maximizers.

43. Thus, since the <u>entire</u> output of a given firm could be taken off the market, and not materially affect market price, the industry's contribution to consumer surplus can be envisaged as coming from the inframarginal firms.

CHAPTER IV: A MEASURE FOR ASSESSING THE WELFARE COST OF A NON-OPTIMAL EXHAUSTIBLE RESOURCE ALLOCATION

A. Introduction and Statement of Purpose

Contrary to the optimistic predictions of the theory of competitive equilibrium, economists realize that the free market does not always direct resources toward their most highly-valued end uses. Although there are several sources of this "market failure," its end result is the same: The total value of goods produced falls short of its potential. The realization that there can be such a waste has prompted economists to develop measures which quantify the degree to which the market fails. Such measures, generally called "welfare-change" indicators, measure the net change in the value output attributable to a given change in the allocation of resources.[1]

In recent years, applied welfare economists have made several contributions in this area of welfare-change measurement.[2] Their studies have generally attempted to develop a measure of the flow of welfare change attributable to a change in a particular set of distortions. The need for such a welfare-change measure in flow markets is necessary in order to rank alternative allocations.

When a government considers intervening in the private market, whether it be in the form of public good provision or direct taxation, it is presumably doing so in order to correct a market failure. It can hardly be argued that a particular sector is a justifiable target for public intervention unless there is some available measure of potential welfare improvement that is both operational and theoretically sound.

Given this interest in determining the scope for a potential Pareto improvement in normal flow market settings, it is especially important to determine if society uses its exhaustible resources at the optimum rate. Furthermore, it would be useful if a measure could be developed which calculated the extent of deviations from that optimum. One reason for concern over the appropriate use rate of an exhaustible resource lies in the importance of properly choosing a discount rate and the social cost of failing to do so: As Solow puts it:

> "It turns out that the choice of a rate of time preference is even more critical in this situation than it is in the older literature on optimal capital accumulation without any exhaustible resources. In that theory, the criterion usually adopted is the maximization of a discounted sum of one-period social welfare indicators, depending on consumption per head, and summed over all time from now to the infinite future. The typical result, depending somewhat on the particular assumptions made, is that consumption per head rises through time to a constant plateau defined by the 'modified Golden Rule.' In that ultimate steady state, consumption per head is lower the higher is the social rate of discount; and correspondingly, the path to the steady state is characterized by less saving and more interim consumption, the higher the social rate of discount. . . When one adds exhaustible resources to the picture, the social rate of time preference can play a similar, but even more critical, role. As a paper by Geoffrey Heal and Partha Dasgupta and one of my own show, it is possible that the optimal path with a positive discount rate should lead to consumption per head going asymptotically to zero, whereas a zero discount rate leads to perpetually rising consumption per head. In other words, even when the technology and the resource base could permit a plateau level of consumption per head, or even a rising standard of living, positive social time preference might in effect lead society to prefer eventual extinction, given the drag exercised by exhaustible resources."[3]

Clearly, the choice of the appropriate discount rate, and the costs of choosing the incorrect one are critical when applied to exhaustible resources. Solow goes on to say:

> ". . . that if exhaustible resources really matter, then the balance between present and future is more delicate than we are accustomed to think; and then the choice of a discount rate can be pretty important and one ought not to be too casual about it."[4]

The purpose, then, of this chapter is to develop a theoretically sound measure for assessing the welfare cost of a suboptimal intertemporal allocation of an exhaustible resource. Although welfare-change measures are not new to the field of welfare economics, such measures have not been developed for exhaustible resources. Therefore, this chapter's welfare-change measure is really an extension of the current work on welfare-change quantification in normal flow market settings.

Before proceeding directly to the welfare-change measure itself, two preliminary subjects must be reviewed. First, before a theoretically sound welfare-change measure can be developed that is applicable to an exhaustible resource, it is essential to briefly review the micro-theoretic underpinnings of welfare-change measures relevant to normal flow market settings. Second, once these underpinnings have been reviewed, the potential sources of welfare change in the market for exhaustible resources must be examined, for it is the <u>net</u> welfare change attributable to these distortions with which this chapter's measure is concerned.

<u>B. A Review of the Modern Welfare Underpinnings of Benefit-Cost Theory</u>

Many of the writings on benefit-cost analysis in recent years[5] have shown that welfare theory is not only relevant, but very much needed when one must evaluate the degree to which "welfare" can be enhanced by a reallocation of resources. Specifically, there is a need to develop a measure of the net benefits of various resource reallocations, in order that they can be ranked among one another on efficiency grounds.

Aside from the unresolved issue of the relative social desirability of alternative income distributions, applied welfare economics has been used in two rather distinct areas.

First, there is the broad area of efficiency benefits. Efficiency benefits are those benefits accruing to society which are attributable to a reshuffling of resources among activities.

Second, welfare economics has also been applied as a tool to evaluate the net benefits of a policy variable <u>which itself</u> alters the resources available to the economy or the technological possibilities under which it operates. Generally, this is not the same as a mere reshuffling of resources among different end-uses.

In short, applied welfare economics can evaluate (1) the net benefits attributable to a reshuffling of resources within a given production possibilities (efficiency benefits), or (2) the benefits of a shift in the production possibilities itself.[6] This dissertation, a study of the allocation of a given stock of an exhaustible resource over time, concentrates exclusively on the determination of efficiency benefits, i.e. the net valuation attribuatble to any given reshuffling of resources among end uses.

In order that there be meaning to "net benefits attributable to a resource reallocation," one must understand that which is commonly accepted by economists to be an efficient allocation in the first place. To quote Currie, Murphy, and Schmitz:

> "Implicit in most of the studies which attempt to identify and measure the welfare effects of resource misallocation is the traditional belief that the competitive equilibrium represents an optimum. As Samuelson observed, 'at least from the time of the physiocrats and Adam Smith there has never been absent from the main body of economic literature the feeling that in some sense perfect competition represented an optimal solution.'"[7]

The authors go on to say:

> ". . . the concept of economic surplus is not necessary, nor has it been found useful, for defining an optimum. In contrast, it has been considered by many to be particularly useful for measuring the welfare effects of deviations from an optimum. It is to this purpose that the majority of its applications are directed."[8]

The above quote seems to imply that economists view efficiency and competition as being somehow synonomous. This is misleading, if not circular.

A Pareto Efficient Allocation

In order that a commodity have an optimum amount of resources devoted to its production in a free-market setting, there are three necessary prerequisites: First, producers must accurately perceive the marginal valuation of the last unit placed on the market, usually, but not always equal to market price. Second, producers must accurately perceive the real marginal opportunity cost on an incremental unit of the good's output, usually, but not always equal to marginal private cost. Third, production must be carried to the point where the two are equal.

Marginal valuation is the maximum amount which some final buyer would be prepared to pay to consume the last unit produced rather than go without. If a given flow of output is offered to the market, and everyone who has a positive preference for it knows of its existence, and each consumer must pay for each unit he consumes and would pay up to his maximum marginal valuation of the last unit rather than go without, the market price will settle at the maximum marginal valuation of the last unit flowing to the market, _if_ buyers are allowed to freely bid for the available quantity. In such a case, since market price is the same as the marginal valuation of the last unit, the producer (if there is one in existence or potentially in existence) will accurately perceive marginal valuation.

Marginal opportunity cost is the value of "other goods" given up by producing an extra unit of the good in question.[9] As an extra unit of a good is produced (in a full employment or "equal unemployment" setting) resources are drawn out of other goods which compete for those resources. If it is true that each resource is hired to a point such that the value of its marginal product equals its price, then by drawing an incremental bundle-of-resources, priced at X dollars, out of alternative employment, there is X dollars of other goods' marginal value lost to the economy. However, as long as the employer of that factor bundle must pay a price of those resources equal to the entire value of those other goods given up, then the employer's marginal private factor costs will be equal to the true marginal opportunity cost, i.e. the marginal value of other goods foregone. If the employer pays any more or less for than bundle, then there is said to be an externality[10] present.

If marginal valuation is accurately reflected in the market price and marginal opportunity cost is equal to marginal private cost, then production will be carried to the point where the two are equal _if_ there is competition in the commodity market and _if_ the bottom of each firm's long-run-average-cost occurs at an output that is small relative to market demand. That is, if there are sufficiently low barriers to entry, such that there is free economic entry and exit of resources, one can expect industry output to be carried to the point where market price equals the marginal private costs of production.

Therefore, if the above three prerequisites are satisfied, the market will allocate resources to the point where the "optimal" output is reached--optimal in the sense that marginal net valuation cannot be increased by any alternative allocation of resources; This is what economists mean by a Pareto efficient allocation of resources toward the production of a given commodity.

It is with reference to this notion of Pareto optimality that measures of welfare change exist. Naturally, whenever the proper amount of resources has already been allocated to the production of a certain commodity, Pareto optimality exists, and an accurate measure of the welfare effect of a resource reallocation would register a zero or negative sign. Conversely, before a welfare-change measure registers positive net benefits, there must be a market failure in the first place. That is, there must be something inherent in a given market which causes a divergence between marginal valuation and marginal opportunity cost.

Careful distinction must be made at the outset between commodity price changes due to (1) a change in the <u>difference</u> between a <u>given</u> marginal opportunity cost schedule and marginal valuation schedule as opposed to (2) a price change resulting from shifts in the schedules themselves. Since this chapter's goal is the development of a measure of efficiency benefits, attention will be directed exclusively to the former.[11] That is, our goal is to assess the net benefits of a partial or total removal of one or more distortions, i.e. divergences between marginal valuation and marginal opportunity costs.

Assumptions

Following a practice which has been traditional in the literature,[12] it will be assumed that all distortions take the form of excise taxes. Also, following Harberger,[13] constant costs will be assumed in the relevant region of the production possibilities. Thus, if there is a change in commodity A's excise tax, the partial equilibrium effect is fully reflected as an equal change in A's consumer price.

The approach taken in this review is to begin the analysis at the individual preference level. After discussing efficiency benefits at the individual level, generalization is made to the market. Next, opportunity costs are introduced, and finally the potential for neutral income transfers are evaluated in order to obtain a net efficiency benefit measure.

Individual Benefits

A measure of aggregate surplus for a single individual can be derived as follows: Suppose there are n commodities, X_i with consumer prices, p_i. Write the consumer's ordinal utility function as follows:

$$\text{IV-1} \qquad U = U(X_1, \ldots X_n)$$

Totally differentiating IV-1, it is seen that increments in utility come about through increments in consumption levels of each of the n goods.

IV-2 $$dU = \sum U_i dX_i$$

where U_i is the marginal utility of the ith good.

A government policy causing a discrete change in the X_i[14] generates the following utility change:

IV-3 $$\int dU = \Delta U = \int \sum U_i dX_i$$

The integral of IV-3 is evaluated from the initial consumption bundle of the X_i's to the new bundle. IV-3, being expressed in utils, is <u>not</u> invariant to a monotonic transformation of the utility function. However, one can calculate a "money income equivalent" of the utility change by dividing the utility differential, dU, by the marginal utility of money income prevailing at each point. That is, as Harberger puts it:

> ". . . by transforming utility into money (income) continuously through the integration process, always at the marginal utility of money prevailing at that point, . . ."[15]

the following measure is obtained:

IV-4 Money income equivalent of utility change = dU/Z

where Z is the marginal utility of money income.[16] This money income equivalent <u>is</u> invariant to a monotonic transformation of the utility index.

Assuming that the consumer is in equilibrium, the following relationship must hold:

IV-5 $$U_i = Zp_i$$

where p_i is the price of the ith good. Substituting IV-5 and IV-2 into the expression IV-4, the following measure is obtained for the money income equivalent of the utility change:

$$\text{IV-6} \qquad \int dU/Z = \int \sum p_i dX_i$$

Equation IV-6 is a measure of the desired money income equivalent of the total utility change attributable to a change in the X_i consumption levels.

Although equation IV-6 is an accurate measure, another measure of the identical income equivalent can be obtained as follows: Calculate the total differential of the consumer's budget equation. The result is equation IV-7.

$$\text{IV-7} \qquad dY = \sum X_i dp_i + \sum p_i dX_i$$

where dY is the differential of the consumer's money income.

Substituting IV-7 into IV-6 yields:

$$\text{IV-8} \qquad \int dU/Z = \Delta Y - \int \sum X_i dp_i$$

where it is to be remembered, that since the consumer is in equilibrium X_i represents the consumer's demand function for each of the i commodities,[17] where X_i is a function of all commodity prices; $X_i = X_i(p_1, \ldots p_n)$.

Both equations IV-6 and IV-8 are identical measures for the utility change brought about by a discrete change in the X_i. However, as will be shown below, equation IV-8 bears a more intuitive relation to the notion of compensating variation, since it measures areas to the left of demand curves. For this reason, the reader's attention is directed to equation IV-8.[18]

Note that equation IV-8, which measures the money income equivalent of the utility gain resulting from a movement from an initial set of prices and money income to a terminal set, has been broken down into two measures. ΔY is an actual money income change, while the term $-\int\sum X_i dp_i$ is a measure dependent on a series of price changes.

Unfortunately, the second term is a line integral, whose value is unique <u>only</u> when $\partial X_i / \partial p_j$ is exactly equal to $\partial X_j / \partial p_i$ for all commodities i and j.[19] That is, the cross effect on p_i on the consumption of X_j must exactly equal the cross effect of p_j on the consumption of X_j. This symmetrical cross effect property, in general, is <u>not</u> satisfied for <u>ordinary demand functions</u>. That is, for example, if the price of electricity falls, whereupon the price of gas falls, the value of the integral will be different than if the order is reversed.

However, as was shown by Hicks and Allen in 1934,[20] these cross terms <u>are</u> equal when all demand curves are income-compensated. Thus, the value of the line integral is unique when adjustments are made in the consumer's money income such that he remains on his initial indifference curve after prices change.[21]

Not only is the value of the line integral unique when such adjustments are made, but specifically, in the case of one discrete (non-marginal) price change, Hicks and Patinkin have shown that the area between the two prices to the left of the compensated demand curve is <u>exactly</u> equal to that unique value of the integral. Hicks chose to call that value the "compensating variation," equal to the compensating variation in money income that would make the consumer indifferent between the initial price and the terminal price. Alternatively stated, it is the money income which must be taken away (given) to keep the consumer on the same indifference curve that he was on when the initial higher (lower) price prevailed.[23]

There are three reasons why the compensating variation is preferable to other measures of welfare gain. First, it bears an intuitive relation to the well-known compensation principle. It is the maximum money income which could be taken away from a gainer after a price change (the minimum which must be paid to a loser after an unfavorable price change) such that he is no better off than before. Second, it is exactly equal to the area left of the income-compensated demand curve,[24] and for that reason, has desirable algebraic-geometric properties. Third, as will be shown below, it is intellectually and theoretically satisfying, since it gives a unique and exact value for the integral in equation IV-8, even when <u>several</u> prices are changing. Hence, future remarks about the compensating variation refer to the integral with the <u>unique</u> value of:

$$\text{IV-9} \qquad C.V. = -\int_{\overline{U}} \sum X_i dp_i$$

where $\overline{U}$ means that the integration process of IV-9 is accomplished in such a way as to keep the consumer at equally preferred levels before and after the price changes.

Aggregation over Commodities

In the case of a single individual, Hicks proved that the value of the line integral of IV-9 is unique even when several prices are changing, as long as the demand curves are income-compensated.[25] However, in addition, Hicks also assured us that the line integral is identically equal to the summed areas under all the compensated demand curves of goods whose prices have changed.[26,27] Hicks outlined a procedure for determining the compensating variation in the case of two related goods. Simply sum the compensating variation associated with the change in the price of X, assuming that the price of Y is fixed at its initial level, and the compensating variation associated with the change in the price of Y, assuming that the price of X is held at its new level.

Hicks does make an important qualification that is especially relevant for public policy, by noting that the more commodities whose prices are changing, the less likely are the income effects to be negligible, and the greater the likelihood of errors from using the relevant areas below ordinary demand curves.[28] This, of course, implies that when demand functions are not income-compensated, the line integral is unique only if the income elasticities are zero for the commodities whose prices are changing.[29] The relevance of Hicks' remark can be most appreciated when it is realized that since not all goods can have a "small" income elasticity, as more goods come under consideration, there is a greater likelihood that estimated consumer surplus to the left of ordinary demand curves will substantially overstate the compensating variation.

Aggregation over Consumers

Before welfare measures are relevant for public policy, one must be able to generalize from the individual to the market. Again, Hicks has determined a solution.[30] Defining the aggregate compensated demand curve as the horizontal summation of individuals' demand curves, Hicks proved that the amount of money income which each consumer would have to lose in order to make each of them as badly off as prior to the price decline, is exactly equal to the area to the left of the relevant aggregate compensated demand curve.[31] This means that the area under the market compensated demand curve is the same as the sum of all the areas under the individual compensated demand curves. This holds for both a single market and for several interrelated markets.

From the standpoint of practical economic policy, therefore, one can attach specific operational meaning to the areas left of ordinary market demand curves. If most consumers have small income elasticities,[32] then the area to the left of an ordinary market demand curve is a fairly accurate measure of the true compensating variation.[33] Furthermore, if prices of related commodities are sequentially or simultaneously altered, areas under all such ordinary market demand curves can be added up as an approximation of the compensating variation accruing to all the relevant consumers, provided that income elasticities are small.[34]

Price Changes, Income Transfers, and Efficiency Benefits

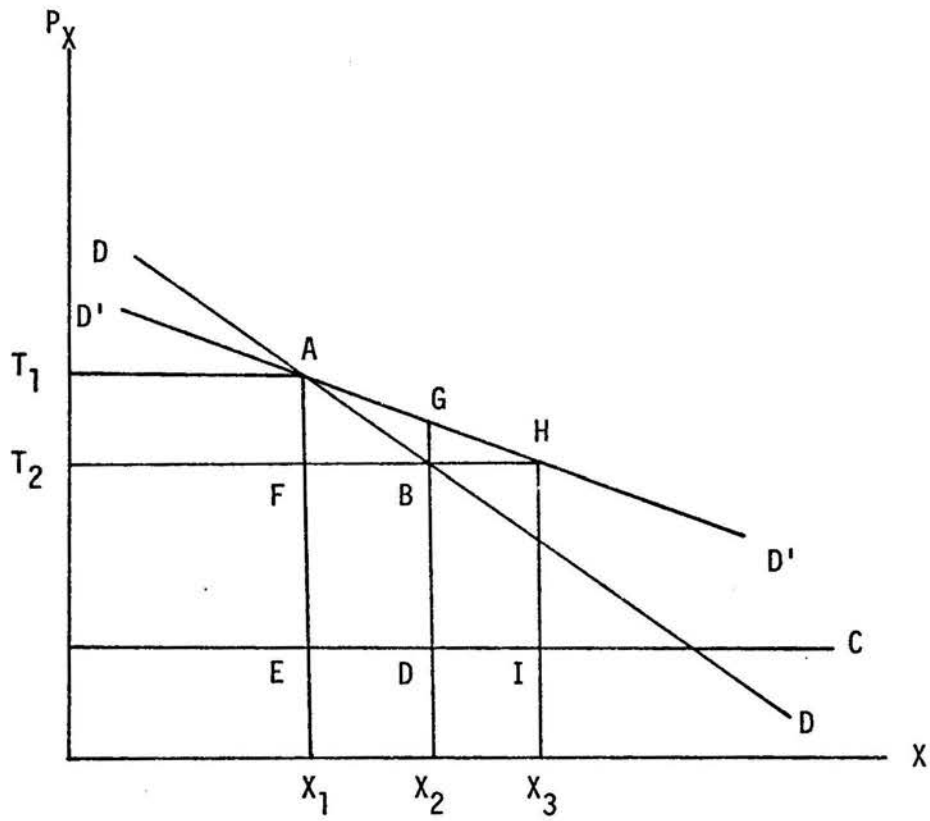
Up to this point it has been shown that the money income change necessary to make the consumer no better off after a price decline is exactly equal to the compensating variation. It has also been argued that the compensating variation is only slightly overstated by the areas left of several ordinary demand curves when income elasticities are small. This notion has been generalized from individual demands to market demands, and it has been demonstrated that areas under market-compensated demand curves are exactly equal to the aggregate of compensating variations of all participants in the market. Furthermore, the summed areas to the left of several ordinary market demand curves only slightly overstate the aggregate compensating variations of all consumers, when (1) the relevant goods have small income elasticities, or (2) form a small part of overall budgets for most of the market participants.

With these relationships in mind, let us bring in the redistributive effects of a price change brought about by a distortion alteration.[35]

Referring to Figure I, let us analyze the net welfare effects, in a general equilibrium setting, of a lowering of the excise tax on good X from T_1 to T_2. Letting DD be the compensated market demand curve and CC be the marginal opportunity cost, it is seen that the tax change lowers the price facing the consumer and hence consumers receive a compensating variation equal to T_1ABT_2.

However, viewing the taxing authority as non-revenue generating,[36] it is seen that the changed tax revenues are matched by an equal but opposite income change.[37] In this case, consumer incomes have been supplemented by an amount equal to BDEF minus T_1AFT_2. Clearly, the resultant net effect of the tax decrease is a net benefit equal to EABD. That is, to obtain the <u>net</u> benefit (money income

FIGURE IV-1



equivalent) of the distortion alteration, in addition to determining the compensating variation in income, one must add the net change in consumer income which would be forthcoming from a non-revenue-raising tax authority.

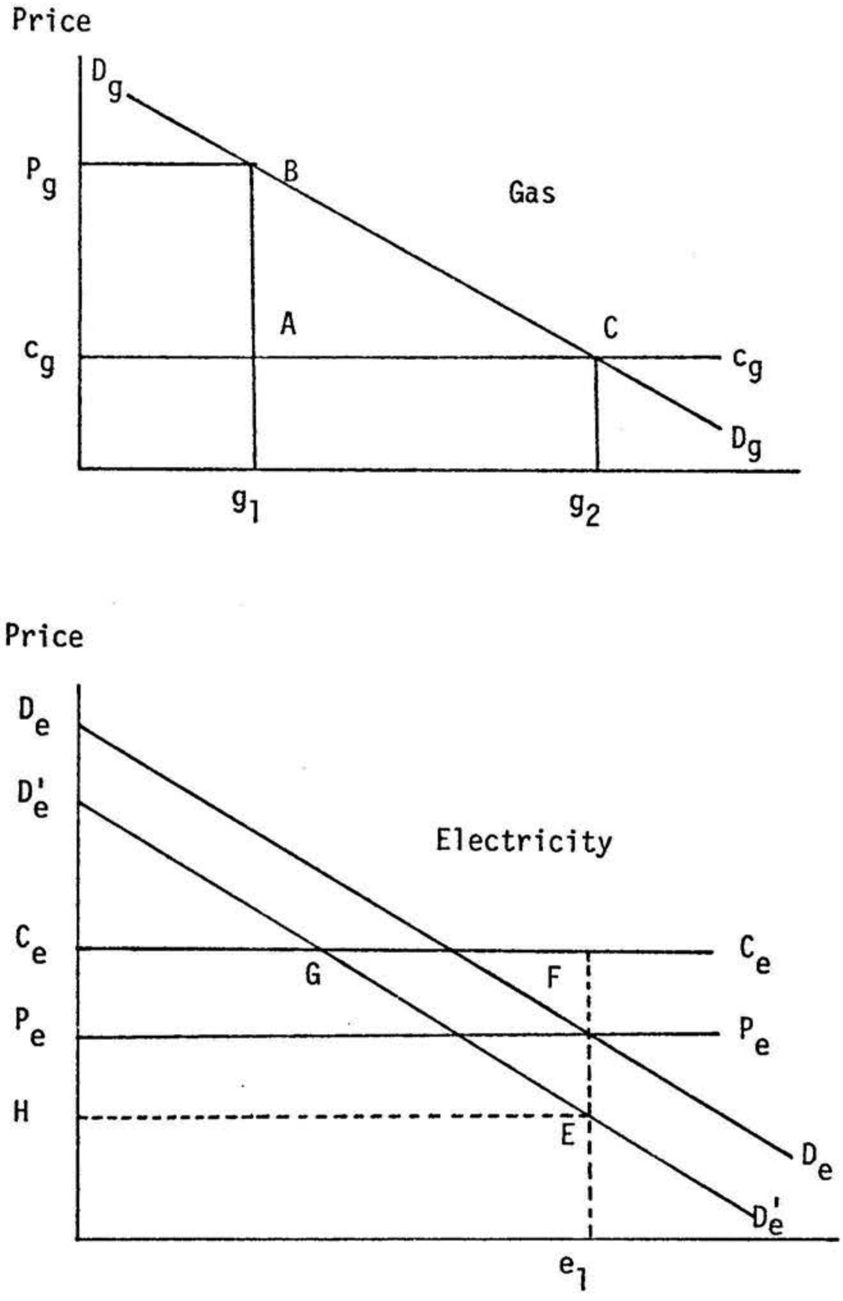
The resultant figure, the trapezoid between the demand and the opportunity cost schedule,[38] is an exact value for the "money income equivalent" of the distortion alteration in a given market. Of course, if distortions are altered in several markets, then the relevant trapezoids under all the respective compensated demand curves must be added together to obtain the aggregate income equivalent of the price changes.[39]

For example, let us assume that two interrelated goods, electricity and gas, are currently being produced in a given region.[40] The gas producer charges a price, p_g, greater than marginal opportunity cost, and produces an output g_1, shown in Figure IV-2. Electricity production, although carried on competitively, emits noxious fumes, thereby imposing real costs in the form of a negative uncompensated externality. Electricity opportunity cost is C_eC_e, greater than private marginal cost, p_ep_e. Given the market demand curve, D_eD_e, electricity production is carried to the point e_1.[41]

To calculate the net efficiency gains of removing both distortions, one can examine the two markets in either order desired, and be certain to exactly account for the money income equivalent of both distortion removals.

Arbitrarily, let us begin with the gas market. Assume that a means is devised, whether it be through some publicly-provided gas, or

FIGURE IV-2


Price
D_g
P_g
B
Gas
A
C
c_g
c_g
D_g
g_1
g_2
Price
D_e
D'_e
Electricity
C_e
G
F
C_e
P_e
P_e
H
E
D_e
D'_e
e_1

through a profit-neutral subsidy, to encourage output to expand to g_2. This increased provision of gas at the correspondingly lower price of c_g will shift the compensated demand curve for electricity to the left, let us say to $D_e'D_e'$. If both demand schedules are income-compensated, an exact measure of the income equivalent of the distortion alteration,

$$\text{IV-10} \qquad \Delta Y - \int_{\overline{U}} \sum X_i dp_i$$

summed over all consumers, is BAC plus EFG,[42,43] equal to the compensating variations, P_gBAc_g minus $HEGc_e$, plus the income transfer of a tax-neutral authority, $HEFc_e$ minus P_gBAc_g.

The measure of net benefits, BAC plus EFT is valid as long as there are no distortions in related markets However, the generalized validity of this benefit-estimation technique is upheld even when there are distortions in markets from which our resources are drawn--as long as the marginal opportunity cost schedule is interpreted carefully. If goods in resource-competing markets are subject to distortions, let us say, taking the form of an excise tax, then the marginal opportunity cost of our good is obtained by adding to its marginal private production costs an amount equal to the tax, since the value of goods given up in other markets exceeds the marginal private cost of the bundle of factors diverted into the production of our good.[44,45]

Summary

This review of the comparative statics of applied welfare economics has shown how one can accurately measure the efficiency benefits attributable to a non-marginal change in a series of distortions, in a general equilibrium setting, within the confines of a given production possibilities and set of consumer tastes.

It should be noted that this review is not intended to be set forth as a unique contribution to the literature. Rather it is the first of two reviews which are deemed necessary as a prerequisite to the development of a net benefits measure attributable to a reallocation of an exhaustible resource.

It is to the second of these two reviews which attention is now directed.

C. Potential Sources of Market Failure in the Market for Exhaustible Resources

There are two potential sources of welfare loss in the market for exhaustible resources. The first source of difficulty lies in the inherent instability of the market. The second source of welfare loss exists even if the market tracks an equilibrium very well. Namely the market could be reacting to the wrong set of signals. Let us discuss each of the sources for market failure in turn.

In chapter two, it was shown that, ideally an exhaustible resource in fixed supply will track a marginal profits path that grows at the rate of compound interest, such that at the time that the last ounce of the resource is taken from the earth, market price will have risen to the point where market demand is just choked off to zero. But there are several reasons why, especially in the market for exhaustible resources, the production path is likely to veer off its equilibrium course.

A major source of disequilibrium in the exhaustive resources market lies in the instability of expectations. Defining an equilibrium price path as being one for which marginal profits grow at the rate of compound interest, and for which the last unit produced is priced at the point where demand is only sufficient to take that unit off the market, then if either marginal profits grow at the wrong rate or if the last unit produced is sold at less than the maximum marginal willingness-to-pay, a disequilibrium price path exists. If, for example, producers think that future marginal profits are increasing faster than the rate of interest, it pays to reduce current production rates (possibly to zero) because they think that the marginal value of a unit in the ground is appreciating faster than the net return from reinvestment of the proceeds from selling the extracted resource.[46]

Conversely, if marginal profits are thought to be rising at a rate less than the rate of interest, producers will increase their extraction rates, rather than leaving in the ground that which is appreciating too slowly. Both of these cases are destabilizing. This is especially true for the latter, since a perceived marginal profit which is rising too slowly, will depress current prices even more, which could, in turn accentuate the excessive current dumping. Eventually mineowners might realize that the mineral was running out and an equilibrium path might be re-established, but the damage would have already been done.

Clearly, what is needed to avoid this social waste is a set of well-developed, long term futures markets.[47] If nothing else, the existence of futures markets would free producers from overreacting to short-term price fluctuations and allow them to make output decisions on the basis of longer-term information.

But even if well-developed futures markets were introduced, the exhaustive resource industry is particularly sensitive to external shocks and uncertainty. Not only is the market likely to fluctuate greatly due to destabilizing price expectations but the uncertainty of future costs, technology, and politics plays a major role. For example if the mineowner is afraid that a substitute technology will soon become available, he might dump excessive current output on the market while there are still some profits to be had. Political uncertainty, although probably not a major factor in the United States, plays a large role in the output decisions of resource owners in foreign countries. Surely if a mineowner thought that his mineral holdings were on the verge of being nationalized, he would not adhere to a self-imposed rationing plan that allowed marginal profits to rise at the rate of interest. His time horizon would be far too short for the overall benefit of society.

Solow, however puts forth a convincing argument that "destabilizing" expectations are not likely to lead to as much instability as the above might imply. Solow emphasizes that an exhaustible resource, in addition to having many normal flow-market attributes, also has characteristics of a capital asset. Specifically, a mineowner perceives his ore as a capital asset which must earn a rate of return equal to the return on other assets of a similar risk. Whereas, in general, the return on an asset is partly dividend and partly capital gain, exhaustive resources, in their natural state, do not earn a dividend, as they do not contribute to production.[48] Therefore the totality of the owner's return on his asset lies in the appreciation of his ore.

Provided that resource owners believe that their resource has a value fixed somewhere in the future, a value determined by technological and demand considerations, then if the current price appears to be rising too slowly toward that point, owners will simply take capital losses on existing stocks, rather than dumping excessive quantities on the market. Thus, as Solow puts it:

> "As well as being destabilized by flow reactions, the market can be stabilized by capitalization reactions. In fact the two stories can be made to merge; the reduction in flow price coming from increased current production can be read as a signal and capitalized into losses on asset values, after which near equilibrium is reached."[49]

So perhaps the resource markets can track their equilibrium paths fairly well.

This brings us to the <u>second</u> source of welfare loss in the exhaustible resources market; namely the failure of profit maximizers to perceive the true social rate of discount, and/or their failure to perceive all relevant opportunity costs. Let us refer to these market failures as stock distortions and flow distortions respectively.

With respect to stock distortions, the literature has argued on two grounds that the private discount rate might be systematically upward biased. On the first ground it has been suggested that the free market will discount its investment decisions at a rate which exceeds the true marginal rate of time preference.[50] Whereas the consuming public will put off consuming a marginal dollar at an interest rate r, the investor must earn a before-tax return of 2r on the marginal dollar invested if he is to earn an after-tax return of r.[51] Clearly, society suffers from an intertemporal misallocation if its marginal investment resources are returning 2r of future consumption while only giving up r worth of current consumption.[52] Thus income taxes on capital which exceed those on consumption cause an upward biased private discount rate.

Another source of bias lies in the role played by risk. Baumol argues that the existence of risk forces private investment to earn something in excess of its expected rate of return.[53] Thus, for example, if the expected rate of return for an investment project of a given risk class is 16 per cent, the private sector might attach a risk premium of 6 per cent to the investment, thus only being willing to pay up to 10 per cent for the use of the resources. Due to the law of large numbers, the 16 per cent return is much more certain from the viewpoint of society. Clearly, the social opportunity cost of a transfer of resources from consumption to investment is 16 per cent and not 10 per cent. Thus <u>some</u> risks for which a premium must be earned on investment resources, are risks to individuals but not to society, and thus lead to too high a private market discount rate.

The second ground on which it has been argued that the discount rate is upward biased is associated with the names of Pigou and Ramsey.[54] According to their argument it is ethically without defense to adopt <u>any</u> discount rate, since by so doing, the present generation would be given preference to future generations. That is, the main reason that we actually do observe positive discount rates is because of the brevity and uncertainty of life. This uncertainty may influence an individual to favor the present over the future, but since unborn generations are every bit as important as the present, they believe, the present should not be given a free rein in the determination of the discount rate. Thus, says the argument, a government which is to be the guardian of future interests as well as current, should not let the present generation's selfishness determine the proper allocation between present and future consumption. Hence, both Pigou and Ramsey suggest a zero discount rate.

Marglin[55] does not accept the Pigou-Ramsey notion of a zero social discount rate. Rather, he "considers it axiomatic that a democratic government reflects only the preferences of the individuals who are presently members of the body politic."[56]

Tullock also rejects the Pigou-Ramsey notion that the social rate should be made less than the private rate. Tullock realizes that such a lowering of the discount rate would increase the redistribution of income from present to future generations, which if past trends are any indication of the future, amounts to taking from the poor and giving to the rich.[57]

Baumol's position would seem to be similar to that of Marglin and Tullock. His conclusion, which he admits is largely a matter of personal opinion, is that

> "by and large, the future can be left to take care of itself. There is no need to lower artificially the social rate of discount in order to increase further the wealth of future generations."[58]

However, Baumol does not rigidly adhere to that opinion. Specifically, he feels that when an irreversible resource allocation is present, as is surely the case with an exhaustible resource,

> "all the wealth and resources of future generations will not suffice to restore them. Investment in the preservation of such items then seems proper, but for this purpose, the appropriate instrument would appear to be a set of selective subsidies rather than a low general discount rate."[59]

Turning to flow distortions, it has been pointed out in the literature that there are at least two such sources that are particularly relevant to the exhaustive resources market. Aside from distortions that are also common to other sectors of the economy,[60] let us direct our attention to two characteristics of the exhaustible resources sector that make it a particularly likely candidate for market failure, namely public goods and open access property rights.

Marglin[61] discusses a "public goods" market failure with respect to the valuation of (current versus future) investment resources, but the argument is equally applicable to exhaustible resources. Since most of us have an uneasy feeling about leaving to future generations a world lacking in some of its natural resources, these resources can be viewed as having the character of a public good. As is the case with other public goods, if each of us knew that everyone else was willing to forego some current use of the exhaustible resource, all of our demand curves would shift to the left. Hence, the sum of all individuals' private demands overstate the true social demand. As Marglin puts it, ". . . the psychic gain from others' investment would outweigh the loss on one's own investment."[62]

The production of natural resources, particularly petroleum resources, is often subject to open access property rights. Both Davidson and Haveman[63] point out that this can lead to a market failure because each of the owners of an open access resource impose negative user costs on one-another. That is, firm A's cost of not mining a parcel of the resource today is not simply its foregone profit on that parcel if mined in the future, but there is also a cost of losing that unit of output to other firms in the area.

The reason that open access property leads to a market failure is that the mutual user cost that each firm imposes on the other is not a real cost from society's perspective. Since the theory of the mine tells us that each producer equates price to marginal cost plus marginal user cost, by each firm increasing output to that point (from its perspective), there is too much production for the social good and the resource will be prematurely exhausted.

The open access distortion is especially acute in the petroleum industry because, not only do each of the firms impose user costs on one another, but excessive current production actually lessens the cumulative extractible petroleum, which is but one more needless cost imposed on society.[64,65]

Although, in principle, the open access distortion can be dealt with by a more careful definition and appropriation of property rights, there are cases, as with an oil pool of unknown size, where property rights are definable at only a high real cost.[66] Given this high cost-of-property-right-definition, perhaps society would suffer less if it tolerated the distortion.

Two flow distortions that are particularly appropriate to the production of exhaustible resources have now been explored. A review of the literature on flow distortions is beyond the scope of this dissertation, but suffice to say, distortions pervade the economy and the production of exhaustive resources is not shielded from their effects.

For the remainder of this chapter, let us direct our attention to the development of a welfare loss measure associated with the second source of market failure, namely the failure of profit-maximizing owners of exhaustible resources to perceive (1) the social rate of discount and (2) all relevant opportunity costs of production.

D. A Model of the Net Welfare Effects of a Time-Reallocation of Mineral Output

In recent years, the literature on applied welfare theory has seen a proliferation of papers which have developed measures of the net welfare effect of various resource reallocations. Such measures are valuable in that they provide an intellectually-satisfying estimate of the net social benefits attributable to those resource reallocations.

The purpose of this section of the chapter is to develop such a measure of the net welfare effect of a reallocation of an exhaustible resource. This measure is fundamentally different from most previously developed "welfare-cost" measures. Whereas most other measures generally ascertain the change in the flow of net benefits attributable to a reallocation of a resource flow, this measure attempts to determine the net time-discounted welfare effect of a change in the time pattern of an exhaustible resource's allocation.

Within the confines of given consumer preferences, a given income distribution, and a given production possibilities set, the allocation of a typical commodity[67] is alterable only by changes in the excise tax[68] placed on that or related commodities. With this in mind one can construct a measure which calculates the welfare effect, on such a typical commodity, of a change in that vector of excise taxes. For example, Harberger, in 1971,[69] developed such a measure. His measure, in differential form,[70] of the welfare effect of a tax, T_j, placed on a single good, is

$$\text{IV-11} \qquad dW = T_j q_j'(T_j)dT_j$$ [71]

where $q_j'(T_j)$ is the marginal effect on q_j's output flow of an excise tax, T_j, and dT_j is the increment to the excise tax, T_j.

At the outset, it should be noted that there are two fundamental differences between the market allocation of a typical commodity and that of an exhaustible resource.

First, as was shown in chapter three, perfect competition in an exhaustible resource market <u>does not</u> lead to an allocation where price equals marginal private cost, as is the case with other goods. Rather, under competitive conditions, the output flow of an exhaustible resource is restricted, at all points in time, to a level <u>short</u> of that where price equals marginal production cost. Each competitive firm <u>restricts</u> its output such that the following <u>private competitive rules</u> are satisfied.

(1) The discounted value of marginal private production costs are equal in all periods.

(2) Average private production costs are minimized in each firm's final period of output.

Second, there are <u>two</u> sources of distortions in the market for an exhaustible resource. In addition to a distortion which upsets the coincidence of private marginal cost and social opportunity cost (the aforementioned "excise tax"), there also exists a possibility that producers might not discount future profits at the same rate that society wishes to discount future surpluses. This "interest rate" distortion is another potential source of the market failing to efficiently allocate an exhaustible resource over time. Let us define these two types of distortions as follows:

$$\text{IV-12(a)} \qquad T_1 = c'_{jp} - c'_j$$

where T_1, the excise tax, is the excess of the ith firm's private marginal cost, c'_{ip}, over its marginal opportunity cost, c'_i. T_1, of course can be of either algebraic sign. The second distortion,

$$\text{IV-12(b)} \qquad T_2 = r - p$$

where the interest rate distortion, T_2, is the difference between the real rate of time preference, r, and the private rate of discount, p.

Following Goldsmith,[72] let us define net discounted welfare over the life of an industry which produces an exhaustible resource:

$$\text{IV-13} \qquad W = \int_{t_o}^{t_1} W(q_1(t), \ldots q_n(t), t)e^{-rt}dt - \sum \lambda_j \int_{t_{o_j}}^{t_{1_j}} q_j(t)dt - K_j$$

where $q_j(t)$ is the time path of the jth firm's extraction of the resource; t_o and t_1 are the beginning and ending period of the industry's operation; t_{o_j}, t_{1_j} are the beginning and ending period of the jth firm's operation; K_j is the size of the jth firm's ownership claim to the resource, and λ_j is a user cost attributable to the jth firm's production, to be discussed further below.

Following Peterson,[73] let us specify Goldsmith's generalized welfare function as consisting of consumer-plus-producer surplus:

$$\text{IV-14} \qquad W(q_1(t), \ldots q_n(t),t) = cs(q(t)) - \sum c_j(q_j(t))$$

where $q(t)$, industry output at a point in time, t, equals $\sum q_j(t)$, the sum of all <u>n</u> firms' output; $cs(q(t))$ is the total area beneath the demand curve[74] up to the level of output, q; and $\sum c_j(q_j(t))$ is the summed value of total (not marginal) opportunity cost across all firms, each producing an output flow, $q_j(t)$.

Given the welfare function specified by IV-13 and IV-14, one can proceed to determine the optimal time path of each of the $q_j(t)$ such that each firm exactly uses up its claim, K_j. One approach to the solution of these optimal $q_j(t)$, using the calculus of variations, was done in chapter three, and has also been done by Goldsmith[75] and others. However <u>another</u> piece of information can be garnered from the welfare measure defined by IV-13 and IV-14.

Suppose that the $q_j(t)$, rather than representing functions <u>to be determined</u> by an efficiency planner, are output paths which have already been determined by the free market, in response to a set of pre-existing distortions, which could be of either the T_1 or T_2 type (equations IV-12). With this interpretation, the W-function of IV-13 is a <u>market-determined</u> welfare function, rather than one which is to be maximized by appropriate choice of the $q_j(t)$. It is very important to note that if the W-function of IV-13 is market-determined by a set of pre-existing distortions, W does not necessarily take on its maximum value. The fact that W is not necessarily being maximized raises the issue of how the λ_j should be interpreted.

When the W-function actually takes on its maximum value, subject to the n resource supply constraints, and only then, will the λ_j be equal to the marginal maximum contribution to welfare attributable to an increment in the stock of the jth firm's resource stock.

If, on the other hand, the value of the W-function is determined by the market's adherence to the previously-mentioned private competitive rules, then the λ_j are not equal to the marginal contribution to welfare of an incremental relaxation of the ith firm's resource stock constraint. In fact, each of the λ_j represent something else. In such a case, each of the λ_j becomes the marginal contribution to privately perceived profits attributable to an addition to the jth firm's resource stock.[76]

Thus, each λ_j represents a shadow price of the resource relative to the function that each individual firm actually maximizes. But, if it so happens that there are no distortions in the market for that exhaustible resource, then the result of each firm maximizing private profits and that of an efficiency planner maximizing the welfare function yield identical shadow prices for the n stocks of the resource. However, such is not the case when the market is subject to pre-existing distortions. Given such a set of market distortions, the market allocates the exhaustible resource on the basis of privately perceived market signals. On the basis of those perceived signals, each of the n firms in the industry allocates its own output over time such that the following holds:

$$\text{IV-15(a)} \qquad [P - c'_{jp}(q_j(t))]e^{-pt} = \lambda_j$$

$$\text{IV-15(b)} \qquad \frac{[P - c_{jp}(q_j(t_{1j}))]e^{-pt_{1j}}}{q_j(t_{1j})} = \lambda_j \qquad 77$$

where P is the resource's price, not perceived as a function of each of the j firms' output, q_j, but which _is_ a function of the collective private output choices over time; $c'_{jp}(q_j(t))$ is private marginal cost of the jth firm's output flow at a point in time, t; p is the private discount rate, not necessarily equal to the real rate of time preference; λ_j is the _private_ shadow price of an increment to the ith firm's resource stock.

Note that equations IV-15(a) and IV-15(b) simply say that each firm should choose an output path such that marginal discounted private profits at each point in time are equal to its constant user cost, and that in the final period of each firm's operation, t_{1j}, average cost should be equal to marginal cost, each being equal to marginal user cost, λ_j.

Remembering that individual firms allocate their stocks of the resource over time on the basis of _private_ costs and _private_ interest, p, it can be said that each firm's time path of output can be written, in general, as follows:

$$\text{IV-16} \qquad q_j(t) = q_j(T_1(a_1,t),T_2(a_2,t),t)$$

where T_1 is the excise tax, a function of time, t, and of a unit parallel shift parameter, a_1, where $T_1'(a_1) = 1$; and where T_2 is the divergence between the real and private rate of discount; T_2 is assumed to be a possible function of time and of a unit parallel shift parameter, a_2, such that $T_2'(a_2) = 1$.

Equation IV-16 says that each firm's output path depends on the time path of the excise tax, the time path of the "interest rate" distortion, and on time itself. Thus, for a given marginal opportunity cost schedule and real rate of time preference relevant to the ith firm, its output path is determined by the time path of (1) the excise tax distortion, and (2) the interest rate distortion.

Assuming that each firm maximizes private profits discounted at the private rate of interest, each will follow the private competitive rules of equation IV-15(a) and IV-15(b); any change in the time path of the excise tax or of the interest rate distortion will cause each firm to choose a new privately optimal time path of output, (although, of course, the socially efficient time paths of output are not affected by the change in distortions for given underlying real cost-valuation schedules). Thus, equation IV-16 summarizes the response of each of the individual firms which follow the private competitive rules of equations IV-15.

In order to construct a measure of the welfare effect of a change in either distortion, T_1 or T_2, one must substitute IV-16, the market response relationship, into IV-14, the welfare measure. After substituting the result into IV-13, the following is obtained:

$$\text{IV-17} \qquad W(a_1,a_2) = \int_{t_o}^{t_1} [cs(\Sigma q_j(T_1(a_1,t),T_2(a_2,t),t)) - \Sigma c_j(q_j(\ .\ ))]e^{-rt}dt - \Sigma\lambda_j \int_{t_{oj}}^{t_{1j}} q_j(\ .\ )dt - K_j$$

where the limits on the first integral refer to opening and closing periods for the industry, and the limits on the integral under the summation sign refer to opening and closing periods for each of the j firms.

With this expression for welfare as an ultimate function of the distortions, a_1 and a_2, differentiation of equation IV-17 gives a measure of how small changes in either distortion, da_1 or da_2 affect welfare, dW.[78]

The calculation of the welfare-loss measure requires differentiation of equation IV-17 both with respect to a parallel shift in the excise tax distortion, da_1, and a parallel shift in the interest rate distortion, da_2. Keeping in mind that the size of a_1 and a_2 affect the opening and closing periods of the industry, t_o and t_1, the opening and closing periods of each firm, t_{oj} and t_{1j}, and the value of each firm's user cost, λ_j, then $W'(a_1)$ and $W'(a_2)$ is equal to the following:

$$\text{IV-18} \qquad W'(a_i) = \int_{t_o}^{t_i} [\sum cs'(q)q'(q_j)q_j'(T_i)T_i'(a_i) -$$

$$\sum c_j'(q_j)q_j'(T_i)T_i'(a_i)]dt + W(t_1)e^{-rt_1}\ t_1'(a_i) -$$

$$W(t_o)e^{-rt_o}\ t_o'(a_i) - \lambda_j \sum \int_{t_{oj}}^{t_{1j}} q_j'(T_i)T_i'(a_i)dt +$$

$$q_j(t_{1j})t_{1j}'(a_i) - q_j(t_{oj})t_{oj}'(a_i) - \sum \lambda_j'(a_j) \int_{t_{oj}}^{t_{1j}} q_j dt$$

$$i = 1,2;\ j = 1, \ldots n$$

where the primes indicate partial derivatives; the $t'(a_i)$ refer to the effect of an increment in a_i on the opening and closing periods of exploitation for both the industry and firms; $W(t_1)$, $W(t_o)$ refer to total surplus (producer plus consumer) at the terminal and initial period of the industry's life, respectively; $q_j(t_{1j})$, $q_j(t_{oj})$ refer to the output flow of the jth firm in its final and initial period of operation, respectively.

With respect to IV-18, note the following observations: $cs'(q)$ is market price, P; Also $q'(q_j)$, the effect on <u>industry</u> output of an increment to a single firm's output, equals one; $T_1'(a_1) = T_2'(a_2) = 1$; $\int_{t_{oj}}^{t_{1j}} q_j dt$, the size of the jth firm's resource stock, equals K_j. With these in mind, IV-18 can be simplified somewhat to:

$$\text{IV-19} \quad W'(a_i) = \int_{t_o}^{t_1} \sum q_j'(T_i)[P - c_j'(q_j)]e^{-rt}dt +$$

$$W(t_1)e^{-rt_1}\ t_1'(a_i) - W(t_o)e^{-rt_o}\ t_o'(a_i) - \sum \lambda_j'(a_i)K_j -$$

$$\sum [\int_{t_{oj}}^{t_{1j}} \lambda_j q_j'(T_i)dt + \lambda_j[q_j(t_{1j})t_{1j}'(a_i) -$$

$$q_j(t_{oj})t_{oj}'(a_i)]]$$

$$i = 1,2;\ j = 1, \ldots n$$

At this point, note the terms (1) $\int_{t_o}^{t_1} \sum q_j'(T_i)[P - c_j'(q_j)]e^{-rt}$ and (2) $\sum \int_{t_{oj}}^{t_{1j}} \lambda_j q_j'(T_i)dt$.

Next, recall the definition of the two distortions of particular relevance to the allocation of exhaustible resources,

$$\text{IV-12(a)} \qquad T_1 = c_{jp}' - c_j'$$

$$\text{IV-12(b)} \qquad T_2 = r - p$$

Assuming that each firm follows the <u>private</u> <u>competitive</u> rules summarized by IV-15(a) and IV-15(b); i.e. each firm allocates its output over time such that both of the following are satisfied:

$$\text{IV-15(a)} \qquad [P - c_{jp}']e^{-pt} = \lambda_j$$

$$\text{IV-15(b)} \qquad [P - c_{jp}/q_j]e^{-pt} = \lambda_j$$

Then, from the definitions IV-12(a) and IV-12(b), it can be said that:

$$\text{IV-20} \qquad [P - c_j'(q_j)]e^{-rt}$$

which is the right hand side of term (1) on the previous page, is equal to:

$$\text{IV-21} \qquad [P - c_{jp}' + T_1]e^{-rt} = [P - c_{jp}' + T_1]e^{-T_2 t}e^{-pt}$$

where the right side of IV-21 comes from the definition of IV-12(b).

However, from the private competitive rule of IV-15(a), IV-21 is simply equal to

$$\text{IV-22} \qquad [P - c_{jp}' + T_1]e^{-rt} = \lambda_j + T_1 e^{-rt}$$

Therefore, it is seen that the difference between the terms (1) and (2) on the previous page, is equal to:

$$\text{IV-23} \qquad \int_{t_o}^{t_1} \sum q_j'(T_i)[P - c_j'(q_j)]e^{-rt}dt - \sum \int_{t_{oj}}^{t_{1j}} \lambda_j q_j'(T_i)dt$$

$$i = 1,2; \; j = 1, \ldots n$$

Continuing with the evaluation of IV-19, note that since the life of each firm must lie within the life of the industry as a whole, the limits of integration, t_o to t_1, include all the limits t_{oj} and t_{1j}. With this observation, the welfare loss measure, IV-19 becomes:

$$\text{IV-24} \qquad W'(a_i) = \int_{t_o}^{t_1} [\Sigma q_j'(T_i)[\lambda_j(e^{-T2t} - 1 + T_1 e^{-rt}]]dt$$

$$+ W(t_1)e^{-rt_1} t_1'(a_i) - W(t_o)e^{-rt_o} t_o'(a_i) -$$

$$\Sigma\lambda_j[q_j(t_{1j})t_{1j}'(a_i) - q_j(t_{oj})t_{oj}'(a_i)] - \Sigma\lambda_j'(a_i)K_j$$

$$i = 1,2; \; j = 1, \ldots n$$

Several points relating to equation IV-24 are worthy of interest.

First, note what the welfare loss measure, IV-24, becomes when the exhaustible resource is in such an abundant supply as to be perceived as effectively having no upper bound. In such a case, since there are no private profits to be had by restricting output short of the level where price equals marginal cost, the private profits of withholding an incremental unit off the market now is zero. If this were the perceived situation, each firm's user cost, λ_j, would also be zero, as would each firm's $\lambda_j'(a_i)$. Furthermore, since there would no private benefits of prolonging the opening of firms, all the t_{jo} and t_o itself would take on zero values, as would $t_{jo}'(a_i)$ and $t_o'(a_i)$. Furthermore, since an abundant resource would be produced as if each firm saw no upper bound on its closing period, both the $t_{j1}'(a_i)$ and $t_1'(a_i)$ would be zero. Under these circumstances, IV-24 becomes:

$$\text{IV-25} \qquad W'(a_i) = \int_0^{\infty} q_j(T_i)T_1 e^{-rt}dt$$

Upon inspection IV-25 is seen to be Harberger's measure, in differential form, IV-11, summed and discounted over the life of the industry.

Returning to equation IV-24, it can be said that the total contribution to society's welfare change can be divided into four parts, each with economic significance.

Let us define part A as the following term:

$$A = \int_{t_o}^{t_1} \sum q_j'(T_i)[\lambda_j(e^{-T_2 t} - 1) + T_1 e^{-rt}]dt; \; i = 1,2; \; j = 1, \ldots n$$

Part A is the effect on welfare of an increment in the distortion, T_i, evaluated throughout the entire period of the industry's operation. The term, $\int_{t_o}^{t_1} \sum q_j'(T_i)T_1 e^{-rt}dt$, as mentioned above, is the dynamic analogue of the Harberger measure, in differential form.

The term $\sum \lambda_j(e^{-T_2 t} -1)$, is the extra effect on welfare (at a point in time) attributable to a pre-existing interest rate distortion. As would be expected, the term is zero when there is no interest rate distortion, and it is larger for larger values of λ_j,[79] the jth firm's private user cost. That is, for a given <u>interest rate</u> distortion, T_2, the welfare cost of an increment to the <u>excise tax</u> distortion is larger when the user cost is larger.

The second part of the total contribution to society's welfare change will be referred to as part B, and is defined as follows:

$$B = W(t_1)e^{-rt_1} \, t_1'(a_i) - W(t_o)e^{-rt_o} \, t_o'(a_i); \; i = 1,2$$

It is seen that B is the effect on aggregate welfare of a tax-induced change in the length of the industry's life.

The first term in B, $W(t_1)e^{-rt_1}\ t_1'(a_i)$, is the discounted value of existing net benefits in the final period, t_1, of the industry's life times the marginal distortion-induced effect on the final period in which the industry operates. This would normally be expected to take on a positive value.

The second term in B, $W(t_0)e^{-rt_0}\ t_0'(a_i)$, is the corresponding discounted value of existing net benefits in the initial period, t_0, of the industry's life, times the marginal distortion-induced effect on the initial period in which the industry operates. This term would also be expected to take on positive values.

The second term would, however, be expected to take on larger absolute values than the first term, because total discounted net benefits in the first period would generally be larger than those same benefits in later periods, as the market response is one of a continuous decrease in output flow over time. The difference between the two terms, i.e. the value of B, then, would be expected to be negative.

The third part of the total contribution to society's welfare will be defined as part C, and is the following term:

$$C = -\sum \lambda_j [q_j(t_{1j}) t_{1j}'(a_i) - q_j(t_{0j})\ t_{0j}'(a_i)]$$

$$i = 1,2;\ j = 1, \ . \ . \ . \ n$$

It is seen that C takes into account the effect of changes in the length of each firm's individual life. The first term in C, $\sum \lambda_j q_j(t_{1j}) t_{1j}'(a_i)$, is the marginal (ith) distortion-induced effect on each firm's final period of operation, t_{1j}', times its output in that final period, $q_j(t_{1j})$, valued at the discounted marginal private user cost, λ_j, summed across all firms.[80]

The second term in C, $\sum \lambda_j q_j(t_{0j})t'_{0j}(a_i)$, is the corresponding marginal (ith) distortion-induced effect on each firm's <u>initial</u> period of operation, $t'_{0j}(a_i)$, times its output in that <u>initial</u> period, $q_j(t_{0j})$, valued at the discounted marginal private user cost, summed across all firms.

Due to the fact that each firm's output is greater in its initial period of operation than in its final period,[81] one would expect the second term in C to have a greater absolute value than the first, therefore leading to a positive number for the overall value of C.

Let us define the final term in the welfare-change measure, IV-24, as part D, which is equal to:

$$D = - \sum \lambda_j'(a_i)K_j$$

It is seen that D measures the summed marginal distortion-induced effect on each firm's <u>unit value</u> of mineral holdings, $\lambda_j(a_i)$, times the corresponding size of each firm's mineral holdings, K_j, summed across all firms. Part D represents the change in the private value of mineral holdings attributable to the distortion change. Since an increase in either distortion reduces the private profitability of mineral extraction, part D must be positive.

<u>E. Comments on the Welfare-Change Measure</u>

Equation IV-24 is a measure of the welfare effect of an increment to either of two generalized distortions relevant to the market for exhaustible resources. From an overall policy perspective however, there may be some practical limitations to the use of the measure. First, there are several conceptual limitations of the measure which should be recognized. Second, even to the extent that the measure is conceptually and intellectually satisfying, some difficult empirical problems would have to be resolved before the measure could be generally useful.

Conceptual Difficulties

With respect to the conceptual limitations, it is important to recognize just what the welfare-change measure fails to take into account. Particularly, note should be taken of the simplifying assumptions made in deriving the measure.

Recall, from section C of this chapter, that the welfare loss measure, equation IV-24, was derived on the assumption that an exhaustible resource tracks its equilibrium price-output path fairly well. Therefore, the measure of welfare change does not include the welfare effect attributable to the market failing to track its equilibrium path.[82] Therefore, to the extent that an exhaustible resource market misses its equilibrium price path, the welfare change measure is biased in the positive direction.[83]

Another major source of oversimplification lies in the assumptions relating to cost and supply of the mineral. Remember that the welfare change measure was constructed on the assumptions of (1) a fixed supply of the resource, unaugmentable by discovery, and (2) a cost function unaffected by cumulative output of the resource. That is, a market response mechanism which assumed (1) and (2) was incorporated into the welfare change measure. In reality, one would expect cumulative output to lead to increasing extraction costs, thus providing an incentive for exploration. Also, one would expect increasing exploration efforts to lead to increased finds of the resource. How does this recognition of economic reality affect the simplified welfare change measure?

In fact, the more readily such exploration efforts lead to mineral discoveries, the more nearly the exhaustible resource approximates a normal flow resource, and the less is the welfare-effect of an interest rate distortion.[84] Therefore, to the extent that rising cumulative production costs lead to increased mineral finds, equation IV-24 overstates the social welfare effect of interest rate distortions.

Another conceptual difficulty of the welfare change measure lies in the implicit assumption that the distortion change is small, relative to its initial size. That is, equation IV-24 measures the ratio of the incremental change in welfare to the incremental change in either of two generalized distortions. In order to calculate an approximate total change in welfare, equation IV-24 must be multiplied by the change in the distortion in question. However the resultant figure is only an approximation of the total change in welfare. An exact measure of the welfare change can be obtained by integrating IV-24 between the old and new values of the distortion.[85] However, by performing the required integration, IV-24 becomes a double integral, which is unduly burdensome. It is for this reason that IV-24 is left in the form of a differential approximation. However, it should be kept in mind that IV-24 is only an approximate measure of the total change in welfare.

The final conceptual factor which is a source for potential confusion is the relationship of the market demand curve to the marginal social valuation schedule. Assuming that efficiency benefits[86] are an appropriate measure of economic welfare, one must exercise caution in the interpretation of the marginal social valuation schedule.[87] Recall that this chapter assumes (1) that all flow distortions take the form of excise taxes which drive a wedge between marginal social valuation and marginal opportunity cost, and not (2) that all distortions are on the cost side of the market. Therefore it is important to recognize that marginal social valuation is not necessarily the same as the market demand curve. For example, there may be sources of marginal social value attributable to an exhaustible resource's use which are not captured in the market demand schedule.[88] For this chapter's welfare-change measure to be interpreted correctly, all such divergences between private demand and marginal social valuation should be included in the flow distortion, T_1.

Empirical Difficulties

In principle, the numerical value of IV-24's welfare change measure can be obtained given a knowledge of each firm's private marginal cost function, each firm's mineral supply claim, the present and future course of private market demand, and the net value of both existing distortions[89] and distortion increments.[90,91,92]

Assuming that each firm acts as if to maximize its total discounted private profits, one can generate the price-output path of a competitive mining industry operating under any given set of market distortions. However, the empirical determination of these paths can pose several problems.

First, even if one can assume that the opening period of all firms and the industry is at time zero,[93] there remains the difficulty of determining (1) the exhaustion period of the industry, t_1, the exhaustion time of each firm, t_{1j}, and (2) the effect of a <u>change</u> in the distortion on the exhaustion period of the industry, $t_1'(a_i)$, and of the firms, $t_{1j}'(a_i)$.[94]

Also, one must be able to determine (1) total consumer-plus-producer surplus in the final period, $W(t_1)$; and (2) each firm's output in its final period of operation, $q_j(t_{1j})$. Note, however, that if firms are generally subject to constant or continually rising average costs, it can be argued on theoretical grounds that both (1) and (2) are equal to zero.[95]

There is also the empirically difficult task of calculating the marginal output response of each firm at <u>each point in time</u>, $q_j'(T_i(t))$, and the marginal cost of each firm's output at <u>each point in time</u>, $c_j'(q_j(t))$.[96] These figures are needed in order that the first term in equation IV-24 can be integrated across the industry's lifetime. Even if each firm's marginal cost function is approximately equal, the data requirements needed to perform the integration are still very great.

In addition, values for each firm's marginal depletion user cost, λ_j, changes therein attributable to distortion shifts, $\lambda_j'(a_i)$, and mineral supplies, K_j, must be known for IV-24 to be operational.

Probably, direct questioning of the firms themselves would yield values that are as good as any estimates by an investigator.

Finally, of course, the welfare change measure assumes that one is able to determine the size of existing distortions, T_i. One surely cannot find a value for the welfare-effect of a change in a distortion if the size of existing distortions is unknown.

F. Summary of Chapter Four

The purpose of chapter four has been to construct a theoretically acceptable measure of the net welfare-effect of a market distortion-induced time reallocation of an exhaustible resource. The approach taken in this chapter was to review (1) the currently accepted welfare underpinnings of benefit-cost theory, and (2) some of the potential sources of market failure in the exhaustible resources sector, before proceeding to the measure itself.

The welfare change measure was constructed on the assumption that each firm which owns a deposit of the exhaustible resource chooses an output path over time consistent with the maximization of its privately-perceived profits. In essence, the welfare change measure incorporated both (1) the effect of a market distortion change on the private-profit maximizing output path of each firm, and (2) the effect of this change in firms' output paths on achievable total discounted net benefits. Therefore, the welfare change measure built on both (1) the private profit-maximizing rules of chapter three and (2) a definition of net social benefits put forth by several writers in the field of exhaustible resources.

Finally, chapter four commented on some of the conceptual and empirical limitations of the welfare loss measure. The conceptual and empirical difficulties are as follows:

Three major conceptual limitations arise from the following simplifying assumptions: (1) the resource market tracks its equilibrium path perfectly, (2) production costs are not affected by cumulative extraction, (3) the supply of the exhaustible resource is fixed and exactly known, and (4) the distortion increment is small relative to its existing size.

The main difficulty in empirical implementation of the welfare loss measure is the sheer volume of data required. Some of the required data include (1) the sensitivity of each firm's output time path to an incremental market distortion; (2) the sensitivity of each firm's "final" period of production to a distortion change; and (3) the sensitivity of each firm's private value of a unit of the resource to an incremental distortion.

Notes for Chapter IV:

1. Such "welfare change" measures lie at the heart of benefit-cost theory. Although the benefit-cost approach estimates the benefits to be had from a resource reallocation, it is implicitly presumed that resources were misallocated in the first place. Otherwise, there would be no gains attributable to an alternative allocation.

2. See for example, R.W. Boadway, "The Welfare Foundations of Cost-Benefit Analysis," Economic Journal, LXXXIV (December, 1974), 926-39; A.C. Harberger, "Three Basis Postulates for Applied Welfare Economics: an Interpretive Essay," Journal of Economic Literature, IX (September, 1971), 785-97; H. Mohring, "Alternative Measures of Welfare Gains and Losses," Western Economic Journal, IX (December, 1971), 349-69; J.B. Shoven, "General Equilibrium with Taxes; A Computational Procedure and an Existence Proof," Review of Economic Studies, XL (October, 1973), 475-90.

3. R.M. Solow, "The Economics of Resources or the Resources of Economics," 9-10.

4. R.M. Solow, "The Economics of Resources or the Resources of Economics," 10.

5. See, for example, R. McKean, Efficiency in Government Through Systems Analysis, Wiley, New York, 1958; J. Krutilla and O. Eckstein, Multiple Purpose River Development, Johns Hopkins University Press, Baltimore, 1958.

6. This distinction is admittedly somewhat artificial, since technological change or other shifts in the production possibilities generally require the use of resources.

7. J.M. Currie, J. Murphy, and A. Schmitz, "The Concept of Economic Surplus and its Use in Economic Analysis," Journal of Economic Literature, LXXXI (December, 1971), 741-99.

8. J.M. Currie, J. Murphy, and A. Schmitz, "The Concept of Economic Surplus and its Use in Economic Analysis," 755.

9. Producers probably are not aware of the marginal opportunity cost of an incremental unit of output--at least not by that name.

10. That is, not exactly all opportunity costs are made internal to the firm.

11. Of course, there are measurable benefits to be had if there is a technological advance which lowers the marginal opportunity cost, but that does not concern us in this review.

12. For example, see A. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay."

13. A. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay," 790.

14. Presumably, this would come about through a change in the vector of excise taxes.

15. A. Harberger, "Three Basis Postulates for Applied Welfare Economics: An Interpretive Essay," 788.

16. The variable Z, the marginal utility of money income, is the extra utility achieved by the consumer when he moves between equilibrium consumption points, after his income rises. Z is a function of all prices, money income, and the utility function.

17. For the moment, ignore how the demand function is compensated.

18. This is the formulation used by M. Burns, "A Note on the Concept and Measure of Consumer's Surplus," American Economic Review, LXIII (June, 1973), 335-44; and by E. Silberberg, "Duality and the Many Consumer's Surpluses," American Economic Review, LXII (December, 1972), 942-52.

19. See any advanced calculus textbook; for example see M. Spiegal, Advanced Calculus, Shaum, New York, 1963, 197.

20. J. Hicks and R. Allen, "A Reconsideration of the Theory of Value," Parts I and II, Economica, New Series, I (February, 1934 and May, 1934), 52-76; 196-219.

21. Actually, for only a single price change, the line integral is unique, regardless of how the consumer is subsidized or compensated. This is because the integral in such a case becomes an ordinary integral with no path dependency difficulty. The difficulty in evaluating the line integral comes when more than one commodity's price changes. Samuelson notes that this non-uniqueness arises due to an assymetrical income effect. See P. Samuelson, "Complementarity--An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," Journal of Economic Literature, XII (December, 1974), 1255-89.

22. J. Hicks, A Revision of Demand Theory, Oxford at the Clarendon Press, Glasgow, 1956.

23. Of course, by making the required money income compensation, the compensating variation becomes equal to "the required change in money income necessary to make the income equivalent ov IV-8 equal to zero."

24. Recall that empirical demand curves are not income-compensated; however when an ordinary demand curve has a small income effect, its area only slightly overstates the compensating variation. This subject is brought up again.

25. J. Hicks, *A Revision of Demand Theory*, 81.

26. *Ibid.*, 81.

27. Of course, it must be kept in mind that some of the curves may be shifting.

28. J. Hicks, *A Revision of Demand Theory*, 179.

29. That this is true can be seen by observing the Fundamental Slutzky Equation: $\partial q_i / \partial p_j = [\partial q_i / \partial p_j]_{\bar{U}} - q_j[\partial q_i / \partial y]_{\bar{p}}$. That is, the total effect of q_i attributable to an incremental change in p_j, is equal to (1) the effect that p_j would have on q_j along a given indifference curve, *minus* (2) an amount equal to q_j times the income effect of the price change. Remember, from the previous section, that for the line integral of IV-9 to be unique the cross price effects must be equal. However, the Slutzky equation says that this cannot happen for ordinary demand curves unless both $\partial q_i / \partial y$ and $q_j /\ y$ are zero (assuming positive consumption of both goods).

30. J. Hicks, "The Rehabilitation of Consumer's Surplus," *Review of Economic Studies*, VIII (February 1940-41), 108-16.

31. A sufficient, but not necessary condition for the aggregate compensated demand curve to coincide with the ordinary market demand curve is that each individual have a zero income elasticity of demand for the commodity.

32. From the Slutzky equation, it is seen that this can happen if either the good constitutes a small part of the consumer's budget or if the income effect is small.

33. In a recent article, Robert Willig has shown how to approximate the compensating variation with only a knowledge of the range of income elasticities in the relevant area of the *ordinary* demand curve. Hopefully, more research will be directed to this area. See R. Willig, "Consumer's Surplus without Apology," *American Economic Review*, LXVI (September, 1976), 589-97.

34. One must remember Hicks' admonition that the more commodities whose prices are changing, the larger will be the total affected budgets and the greater will areas to the left of ordinary demand curves be overstatements of the compensating variation.

35. In a constant-cost economy, factor prices do not change in the relevant region. Hence, all redistributive effects come from the demand side rather than the factor price side. This constant cost assumption is made to avoid some of the unresolved difficulties associated with producer's surplus and economic rent.

36. That is, all tax revenues are returned to the consumers as income supplements, and all tax decreases are financed by a relowering of consumer incomes.

37. If, for example, the distortion in question is the lowering of monopoly profits, one can still view the redistribution as a neutral income transfer from one sector (the monopolist) to another (the consumer).

38. This measure has been referred to as "marginal valuation" in the literature. Note that if the output change in question, dX_i, is small, then the marginal valuation of dX_i, becomes a shrunken slice of the trapezoid, and is equal to T_i, the size of the distortion in that market. Furthermore, if dX_i is brought about by a change in T_i, then the marginal benefits attributable to dT_i are equal to $T_i X_i'(T_i)dT_i$. This measure of net benefits, in incremental form, is articulated more fully in A. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay." Also see chapter four of this dissertation for the development of a similar measure relative to the market for an exhaustible resource.

39. If D'D' is the respective ordinary demand curve, then measured benefits would be AHIE, an overstatement of GHID.

40. This example is used by E. Mishan, Cost-Benefit Analysis, Praeger, London, 1972, 40-1.

41. Analytically, one might say that the equivalent of a positive tax, $p_g - c_g$, is being levied on gas consumers, and a negative tax, $p_e - c_e$, on electricity consumers.

42. BAC plus EFG is the appropriate measure if other substitutable goods are produced in distortion-free markets. Otherwise another adjustment must be made. If, for example, the above price changes induced ΔY more expenditures in the coal sector, which was subject to an excise tax of t_c, then a non-revenue-raising tax authority would return $\Delta Y[t_c]$ to these new coal consumers. This income change must also be counted as a benefit, since it is part of the overall ΔY.

43. On this subject, Patinkin remarked that "it would be meaningless to attempt to partition this aggregate measure of welfare change into 'that part due to the change in the price of X and that part due to the change in the price of Y.'" But, as he emphasizes, "it is difficult to conceive of policy questions for which such a disentanglement would be of interest. For what concerns us in practice is the total surplus generated by any particular action. The abstract accounting imputation of this total to particular commodities is of no operational significance." D. Patinkin, "Demand Curves and Consumer's Surplus," 101.

44. Technically, this means that one should interpret such a related market distortion as a negative tax, for which there would be zero net benefits of increasing the output of our good, if there were an equal excise tax in our market.

45. This method of compensating for distortions elsewhere was suggested by Hicks in "The Rehabilitation of Consumer's Surplus." However Mishan, in "A Survey of Welfare Economics, 1939-1959," Economic Journal, LXX (June, 1960), 197-256, argued that this technique breaks down when prices depart from marginal costs in different degrees in different industries. Mishan is undoubtedly correct, and for a proper correction of "our good's" marginal cost, one should attempt to determine the degree to which our bundle of resources is being drawn from distorted-versus-distortion-free sectors.

46. This "reinvestment-of-the-proceeds" forumulation is used by H.W. Richardson, in Economic Aspects of the Energy Crisis, D.C. Heath and Company, New York, 1975, 43.

47. R.M. Solow, "The Economics of Resources or the Resources of Economics," thinks that this is the main area in which public policy will be able to make a substantial contribution to long-run stability. By initiating the development of such markets, government intervention would probably be more beneficial than if it attempted institutional solutions.

48. This assumes that society does not directly value the resource in its natural state. To the extent that there is a willingness-to-pay for a conservation ethic, or for maintaining future options, an idle exhaustible resource does earn a dividend.

49. R.M. Solow, "The Economics of Resources, or the Resources of Economics," 6-7.

50. The marginal rate of time preference is that rate that is necessary to induce consumers to postpone an incremental dollar worth of resources consumed.

51. This assumes an across-the-board corporate income tax of 50 per cent. For an excellent treatment on the subject of the discount rate, see W.J. Baumol, "On the Social Rate of Discount," American Economic Review, LVIII (September, 1968), 788-802.

52. This assumes no marginal income tax on current income consumed.

53. W.J. Baumol, "On the Social Rate of Discount," 796.

54. A.C. Pigou, The Economics of Welfare, 4th edition, Macmillan, London, 1932, 22-30; F.P. Ramsey, "A Mathematical Theory of Saving," Economic Journal, XXXVIII (December, 1928), 543-9.

55. S.A. Marglin, "The Social Rate of Discount and the Optimal Rate of Investment," Quarterly Journal of Economics, LXXVII (February, 1963), 95-111.

56. S.A. Marglin, "The Social Rate of Discount and the Optimal Rate of Investment," 97.

57. G. Tullock, "The Social Rate of Discount and the Optimal Rate of Investment: Comment," Quarterly Journal of Economics, LXXVIII (May, 1964), 331-6.

58. W.J. Baumol, "On the Social Rate of Discount," 801.

59. Ibid., 801.

60. Such distortions are high market concentration, taxation, and technical externalities in both the factor and commodity markets.

61. S.A. Marglin, "The Social Rate of Discount and the Optimal Rate of Investment," 95-111.

62. S.A. Marglin, "The Social Rate of Discount and the Optimal Rate of Investment," 103.

63. R. Haveman, "Common Property, Congestion, and Environmental Pollution," Quarterly Journal of Economics, LXXXVII (May, 1973); P. Davidson, "Public Policy Problems of a Domestic Crude Oil Industry," American Economic Review, LII (March, 1963), 85-108.

64. This point is emphasized by Davidson, "Public Problems of a Domestic Crude Oil Industry."

65. There have been attempts to lessen this source of market failure through government imposition of well-spacing requirements and maximum pumping rates, but unless resource ownership rights are vested in a centralized private body, there is bound to be some arbitrariness in such public standards.

66. For a classic treatment of common-property resource distortions, see H.S. Gordon, "The Economic Theory of a Common-Property Resource: the Fishery," Journal of Political Economy, LXII (April, 1954), 124-42.

67. It is assumed that a "typical commodity" is one whose allocation is unaffected by changes in the private rate of discount. To the extent that a change in the rate of discount affects the allocation of the fixed capital with which the commodity is produced, this is an oversimplification.

68. It is assumed that "excise tax" stands for any flow distortion associated with the production or consumption of a commodity. The "excise tax," therefore includes such distortions as monopoly, externalities, factor market imperfections, and public goods.

69. A. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay."

70. "Differential form" is that form which measures the welfare effect of a small (differential) change in the vector of relevant distortions.

71. Harberger proved that, if distortions in other markets are zero, then IV-11 captures the total welfare effect of such an excise tax change. See section "E" of this chapter.

72. O. Goldsmith, "Market Allocation of Exhaustible Resources."

73. F. Peterson, "The Theory of Exhaustible Natural Resources: A Classical Variational Approach."

74. Ideally, the demand curve should be income-compensated for consumer surplus to be an accurate measure of the gross efficiency benefits of a given level of output flow.

75. O. Goldsmith, "Market Allocation of Exhaustible Resources."

76. This assertion comes directly from the theory of LaGrange multipliers. It can be shown (See M. Intriligator, Mathematical Optimization and Economic Theory, Prentice-Hall, Englewood Cliffs, New Jersey, 1971, 20-43.) that in the case of a single constraint, the λ is equal to the increment in the optimal value of the objective function per unit relaxation of the constraint. That is, the LaGrange multiplier, at the solution, measures the sensitivity of the optimal value of the objective function, $F^* = F(x_1^*, \ldots x_n^*)$, to variations in the constraint constants. When the objective function is not being optimized, the LaGrange multipliers have no such interpretation.

77. These equations come directly from the results of chapter three, equations III-29(a) and III-29(b), and arise from each firm acting as if to maximize the time-discounted value of private profits, equal to $\int_{t_{0j}}^{t_{1j}} [[pq_j - c_{jp}(q_j)]e^{-pt} - \lambda_j(q_j)]dt$.

78. Such differentiation gives an approximation to the change in the function $W = W(a_1,a_2)$. The approximate change, dW, is the total differential of W, i.e. $dW = W'(a_1)da_1 + W'(a_2)da_2$. In this section of the chapter, calculations are made for $W'(a_i)$, $i = 1,2$.

79. Recall, from chapter two, that factors which lead to a larger λ_j are lower private extraction costs, smaller mineral supplies, and a lower private interest rate.

80. Recall that since discounted net benefits are generally not at their maximum when distortions exist, λ_j is not a marginal contribution to net social benefits. Rather, as shown earlier, λ_j is a term reflective of the marginal private profits of an increment to the ith firm's resource stock. One would expect λ_j to be less than the maximum possible contribution to net benefits.

81. See chapter three.

82. From section C of this chapter, the equilibrium path was defined as the one on which (1) the output path forthcoming from the industry gives rise to a price path along which each firm's marginal private profits grow at the rate of private interest, and (2) the last unit of the resource stock mined brings a market price which is exactly at the level where quantity demanded is equal to that one unit. If producers put too much output on the market over a given length of time, such that marginal profits grow at less than the rate of interest, then the resource is being exhausted more quickly than the equilibrium rate. Likewise if the price of the last unit mined is less than the "choke" price on the demand curve, the resource is exhausted at a faster-than-equilibrium-rate.

83. That is, there is an additional social welfare loss attributable to production veering from the equilibrium path.

84. Recall that there is no welfare effect of an interest rate distortion in a normal flow market.

85. That is IV-24 measures $W'(a_i)$ for $i = 1,2$. An <u>approximation</u> to the total change in welfare, dW, is equal to $W'(a_i)da_i$. The exact change in welfare, ΔW, is equal to $\int_{a_{io}}^{a_{i1}} W'(a_i)da_i$; $i = 1,2$.

86. Recall that efficiency benefits have nothing to do with the distribution of income. Rather, they measure the "equivalent" of the increase in aggregate income attributable to a reduction in one or more market distortions.

87. Actually, this schedule does not appear explicitly in the welfare loss measure, but the divergence between it and the marginal opportunity cost schedule, T_1, does appear. Therefore, to understate marginal social valuation is to understate T_1.

88. Three examples are option value, preservation value, and existence value.

89. Note, that according to equation IV-24, the size of the <u>existing</u> distortions must be known in order that total net social valuation (consumer-plus-producer surplus) can be determined in both the beginning and ending period of the industry's life, i.e. $W(t_0)$ and $W(t_1)$.

90. Recall, from section D, that IV-24 measures the welfare effect of a <u>change</u> in either of two market-distortions relevant to the production of an exhaustible resource, (1) a divergence between the market rate of interest and social rate of time preference, or (2) a distortion that causes a divergence between marginal social valuation and market demand or marginal opportunity cost and marginal private cost.

91. Recall that the distortion, T_i, is assumed to be a function of time and a parallel shift parameter, a_i, i.e. $T_i = T_i(a_i,t)$. Therefore $W'(a_i)$ measures the effect on welfare of a <u>parallel</u> shift over time of the distortion, T_i; i.e. the welfare effect of a shift in the distortion by an equal amount at all points in time.

92. Even with the most elementary cost, demand, and distortion functions, the author was unable to calculate an analytical solution for $W'(a_i)$, because some of the variables could not be expressed as explicit functions of other required variables. Therefore, even for a simple set of cost and demand functions, numerical approximation methods are necessary to estimate the value of IV-24.

93. An opening period of time zero can be interpreted as meaning that IV-24 measures the welfare effect of a distortion increment relative to an exhaustible resource which is <u>already</u> being produced.

94. In principle, t_1, t_{1j}, $t_1'(a_i)$, and $t_{1j}'(a_i)$ can be determined by examining each of the n firms' transversality conditions. See "An Example," Chapter III, Model I.

95. See "An Example," in Chapter III, Model I.

96. When a welfare change function is constructed for a normal flow commodity, only an aggregate output response, and an aggregate industry supply price are needed to make the necessary calculations. (See A. Harberger, "Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay.") However, recall that an "industry supply curve" for an exhaustible resource has no meaning. See chapter II, part B.

CHAPTER V: SUMMARY AND CONCLUSIONS

This dissertation has been an attempt to provide a conceptual foundation upon which (1) one can assess the desirability of the free market's time allocation of an exhaustible resource, and (2) one can measure just how well the free market's allocation of the exhaustible resource approaches a Pareto efficient allocation.

In the second chapter (the first chapter after the literature review), a series of three comparative statics models were constructed which showed the implications of changes in the supply of either of two mineral grades, extraction costs, or the discount rate, on the market-determined outcomes of price, unit profits, extraction time, and attainable net benefits. This second chapter was to a large extent, a formalization of the literature on the comparative statics of the competitive mining industry under certainty.

The third chapter consists, first, of an introduction to the mathematics needed in solving those exhaustible resource problems which require the calculus of variations. Next, two pairs of models were developed. In the first two models, a comparison was made between the allocation of an exhaustible resource claim by a hypothetical one-firm price taker, and that same claim when operated by a socialist efficiency planner. The one-firm industry was shown to deplete the resource too quickly. The second pair of models was an *n*-deposit generalization of the first models. Specifically, the output path forthcoming from *n* profit maximizing competitive mines was compared with that of the same *n* deposits operated by an efficiency planner. It was shown that under certain assumptions, the two outcomes are identical. Although this result is not new to the literature, it appears to have been suggested by previous authors on the basis of insufficient proof.

The purpose of the fourth chapter was to develop a theoretically satisfactory measure of the net welfare effect of a distortion-induced time reallocation of an exhaustible resource. The first part of this chapter introduced the notion of welfare loss and described some of the sources of those losses in the exhaustible resource sector, before developing the measure. The measure itself incorporated (1) the effect of a market distortion change on the (private) profit-maximizing output path of each of the _n_ firms, and (2) the effect of this change in each firm's output path on aggregate total discounted net benefits.

In order that this welfare change measure be put into proper perspective, let us turn to policy questions for which the measure has some relevance.

The welfare change measure developed in chapter four is the dynamic analogue of welfare change measures applicable to the market for _non_-exhaustible outputs. Such measures for flow outputs approximate the equivalent loss in real output due to a distortion-induced resource misallocation. Such measures can be used to rank alternative policies on efficiency grounds, and are at the heart of modern cost-benefit analyses which measure the welfare effect of increments to output in distorted markets. Thus, for example, if a welfare change measure indicated net benefits of \$3 million in a project A and \$2 million in a project B of equal costs, then project A would be judged preferable by $1 million on efficiency grounds.

The need for developing such a ranking measure for exhaustible output is clear when one considers the multiplicity of exhaustible resource policy decisions with which the public sector is confronted. For example, the efficiency benefits of a partial breaking up of monopoly power in the domestic coal industry could be evaluated by the use of this measure. That is, assuming that the only divergence between the marginal social valuation and marginal opportunity cost of coal production is the excess of monopoly price over marginal cost, then the amount by which the industry reorganization brought price closer to marginal cost would be equal to the change in the market distortion in chapter four's welfare change measure. An approximation to the net welfare effect of the distortion change would be given by that change times the value of the welfare change measure in chapter four.

From the example given above, it is seen that the scope for use of the welfare change measure is wide, indeed. As noted in the text of the dissertation, the "excise tax" is just a generic term which includes any flow distortion, from whatever source it may originate. In this sense, "excise tax" includes the excess of monopoly price over marginal cost, externalities of all kinds, or any other source of divergence between marginal valuation and marginal opportunity cost.

However, lest one become too enthusiastic with the welfare change measure, the reader should be warned of its shortcomings, for the measure can be easily misused. For example, it would not always be a simple matter to determine if an energy tax on coal production reflected an increase or a decrease in the existing size of the distortion in the coal market. Would the increased tax actually improve the time-allocation of coal by forcing society to internalize more of the opportunity costs of extraction, or would the tax be adding to existing inefficiencies by encouraging a reduction in a current output flow which is already too low? Needless to say, careful economic judgements would have to be made about the algebraic sign of existing distortions before being able to evaluate the welfare effect of alternative policies.

Hopefully, however, the welfare change measure which has been developed in this dissertation will provide a good starting point for more theoretical research. Additional research is needed at both the level of empirical usefulness and at the level of conceptual completeness.

At the level of empirical relevance, research could be fruitfully directed toward the development of a more simplified welfare change measure--one which places less demand on data requirements. It would especially be more desirable from a practical standpoint if output data were not required for all points in time from each firm in the industry.

Relative to the conceptual level, the reader will recall that chapter four's measure was derived on the basis of several particularly simplifying assumptions, two of which point to the need for additional research.

First, the measure assumes that each deposit of the resource is fixed, unaugmentable by discovery effort, and subject to constant cumulative extraction costs. Research is needed in the development of a measure which incorporates the effect of increasing cumulative extraction leading to rising extraction costs with the resulting increased incentive to explore for new deposits.

Second, the measure does not take into account any source of welfare loss attributable to the resource tracking a disequilibrium price-output path. Several authors cited in this dissertation have argued that such an equilibrium path would, in fact, probably be missed in the absence of a set of ideal futures and contingent commodities markets. Since ideal futures and contingent commodities markets do not exist for exhaustible resources, additional research is needed in the development of a measure which incorporates the effect of this destabilizing force as an additional source of welfare change.

However, weaknesses aside, the welfare change measure was constructed on the basis of a received theory of exhaustion, which was extended in chapters two and three of this dissertation. Therefore, in spite of some of the simplifying assumptions with which the theory of exhaustion deals, this dissertation's welfare change measure has attempted to provide a point of departure for which alternative exhaustible resource allocations can be meaningfully compared.

BIBLIOGRAPHY

1. Aldrich, C.A. "A Model for Funding Interrelated Research and Development Projects Under Uncertainty," Journal of Financial and Quantitative Analysis, IX (January, 1974).

2. Banks, F.E. "A Note on Some Theoretical Issues of Resource Depletion," Journal of Economic Theory, IX (October, 1974).

3. Bator, F.M. "The Anatomy of Market Failure," Quarterly Journal of Economics, LXXII (August, 1958).

4. Baumol, W.J. "On the Social Rate of Discount," American Economic Review, LVIII (September, 1968).

5. Bellman, R.E. Dynamic Programming, Princeton, N.J.: Princeton University Press, 1957.

6. Broadway, R.W. "The Welfare Foundations of Cost-Benefit Analysis," Economic Journal, LXXXIV (December, 1974).

7. Burns, M.E. "A Note on the Concept and Measure of Consumer's Surplus," American Economic Review, LXIII (June, 1973).

8. Burt, O. and R. Cummings. "Production and Investment in Natural Resource Industries," American Economic Review, LX (September, 1970).

9. Canarella, G. "Optimal Policies for Depletable Resources," (unpublished Ph.D. dissertation, University of Virginia), 1973.

10. Carlisle, Donald. "The Economics of a Fund Resource with Particular Reference to Mining," American Economic Review, LXIV (September, 1954).

11. Chiang, Alpha. Fundamental Methods of Mathematical Economics, New York: McGraw-Hill, 1974.

12. Ciriacy-Wantrup, S.V. "Taxation and the Conservation of Resources," Quarterly Journal of Economics, LII (February, 1944).

13. Cummings, R.G. "The Economic Theory of Exhaustible Resources: Some Extensions," (unpublished Ph.D. dissertation, University of Kansas), 1968.

14. Cummings, R.G. "Some Extensions of the Economic Theory of Exhaustible Resources," Western Economic Journal, VII (September, 1969).

15. Cummings, R.G. and O. Burt. "The Economics of Production from Natural Resources: Note," *American Economic Review*, LIX (December, 1969).

16. Currie, J.M., J.A. Murphy, and A. Schmitz. "The Concept of Economic Surplus and its Use in Economic Analysis," *Economic Journal*, LXXXI (December, 1971).

17. Dasgupta, P. and G. Heal. "The Optimal Depletion of Exhaustible Resources," *Review of Economic Studies Symposium*, (summer, 1974).

18. Davis, R.M. "The Current State of Profit Theory," *American Economic Review*, XLII (June, 1952).

19. Dorfman, R. and N. Dorfman. *Economics of the Environment*, second edition, New York: W.W. Norton, 1977.

20. Eckstein, O. *Water Resource Development: The Economics of Project Evaluation*, Cambridge, Mass.: Harvard University Press, 1958.

21. Gaffney, M. *Extractive Resources and Taxation*, Madison, Wisconsin: The University of Wisconsin Press, 1967.

22. Goldsmith, O. "Market Allocation of Exhaustive Resources," *Journal of Political Economy*, LXXXII (September-October, 1974).

23. Davidson, P. "Public Policy Problems of the Domestic Crude Oil Industry," *American Economic Review*, LIII (March, 1963).

24. Gordon, H.S. "The Economic Theory of a Common-Property Resource: the Fishery," *Journal of Political Economy*, LXII (April, 1954).

25. Gordon, R.L. "Conservation and the Theory of Exhaustible Resources," *Canadian Journal of Economics*, XXXI (August, 1966).

26. Gordon, R.L. "A Reinterpretation of the Pure Theory of Exhaustion," *Journal of Political Economy*, LXXV (February, 1967).

27. Gray, L.C. "Rent Under the Assumption of Exhaustibility," *Quarterly Journal of Economics*, XXVIII (May, 1914).

28. Harberger, A.C. "Three Basic Postulates for Applied Welfare Economics: an Interpretive Essay," *Journal of Economic Literature*, IX (September, 1971).

29. Haveman, R. "Common Property, Congestion, and Environmental Pollution," *Quarterly Journal of Economics*, LXXXVII, (May, 1973).

30. Heal, G. "Economic Aspects of Natural Resource Depletion," in Economics of Natural Resource Depletion, D.W. Pearce and J. Rose (eds.), London, Macmillan, 1975.

31. Herfindahl, Orris. "Depletion and Economic Theory," in Mason Gaffney, Extractive Resources and Taxation, Madison, Wisconsin: University of Wisconsin Press, 1967.

32. Hicks, J.R. "The Rehabilitation of Consumer's Surplus," Review of Economic Studies, VIII (February, 1940-41).

33. Hicks, J.R. A Revision of Demand Theory, London: Oxford University Press, 1956.

34. Hicks, J.R. and R.G.D. Allen. "A Reconsideration of the Theory of Value," Parts I and II, Economica, New Series, I (February, 1934; May, 1934).

35. Hotelling, H. "The Economics of Exhaustible Resources," Journal of Political Economy, XXXIX (April, 1931).

36. Houthakker, H. and R.M. Solow. "Comments and Discussion on 'The Allocation of Energy Resources,'" Brookings Papers on Economic Activity (3: 1973).

37. Intriligator, M. Mathematical Optimization and Economic Theory, Englewood Cliffs, N.J.: Prentice-Hall, 1971.

38. Jevons, W.S. The Coal Question, London: Macmillan, 1865.

39. Kay, J. and J. Mirrlees. "The Desirability of Natural Resource Depletion," in D. Pearce and J. Rose, Economics of Natural Resource Depletion, London, Macmillan, 1975.

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

41. Krutilla, J. and O. Eckstein. Multiple Purpose River Development, Baltimore: Johns Hopkins University Press, 1958.

42. Lee, D.R. and D. Orr. "The Private Discount Rate and Resource Conservation," Canadian Journal of Economics, VIII (August, 1975).

43. Lipsey, R. and K. Lancaster. "The General Theory of Second Best," Review of Economic Studies, XXIV (February, 1956).

44. Marglin, S.A. "The Social Rate of Discount and the Optimal Rate of Investment," Quarterly Journal of Economics, LXXVII (February, 1963).

45. McDonald, Stephen. "Percentage Depletion, Expensing of Intangibles, and Petroleum Conservation," in Mason Gaffney, Extractive Resources and Taxation, Madison, Wisconsin: University of Wisconsin Press, 1967.

46. McKean, R. Efficiency in Government Through Systems Analysis, New York: Wiley, 1958.

47. Mishan, E.J. "A Survey of Welfare Economics, 1939-1959," Economic Journal, LXX (June, 1960).

48. Mishan, E.J. Cost-Benefit Analysis, New York: Praeger, 1971.

49. Mohring, H. "Alternative Measures of Welfare Gains and Losses," Western Economic Journal, IX (December, 1971).

50. Nordhaus, W. "Theory of Innovation: An Economic Theory of Technical Change," American Economic Review, LIX (May, 1969).

51. Nordhaus, W. "The Allocation of Energy Resources," Brookings Papers on Economic Activity, (3:1973).

52. Nordhaus, W. and J. Tobin. "Is Growth Obsolete?" in National Bureau of Economic Research, Economic Growth, Fiftieth Anniversary Colloquium, Volume V, New York: National Bureau of Economic Research.

53. Patinkin, D. "Demand Curves and Consumer's Surplus," in Carl Christ et al., Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld, Stanford: Stanford University Press, 1963.

54. Pearce, D.W. and J. Rose. Economics of Natural Resource Depletion, London: Macmillan, 1975.

55. Peterson, F.M. "The Theory of Exhaustible Natural Resources: A Classical Variational Approach," (unpublished Ph.D. dissertation, Princeton University), 1972.

56. Pigou, A.C. The Economics of Welfare, London: Macmillan, 1932.

57. Pontryagin, C.S. et al. The Mathematical Theory of Optimal Processes, New York: Interscience Publishers, 1962.

58. Richardson, H.W. Economic Aspects of the Energy Crisis, New York: D.C. Heath and Company, 1975.

59. Ramsey, F.P. "A Mathematical Theory of Saving," Economic Journal, XXXVIII (December, 1928).

60. Salant, S.W. "Exhaustible Resources and Industrial Structure: A Nash Cournot Approach to the World Oil Market," Journal of Political Economy, LXXXIV (October, 1976).

61. Samuelson, P.A. "Complementarity--An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," Journal of Economic Literature, XII (December, 1974).

62. Scott, A.T. "Notes on User Cost," Economic Journal, LXIII (June, 1953).

63. Scott, A.T. "The Theory of the Mine Under Conditions of Certainty," in Mason Gaffney, Extractive Resources and Taxation, Madison, Wisconsin: University of Wisconsin Press, 1967.

64. Silberberg, E. "Duality and the Many Consumer's Surpluses," American Economic Review, LXII (December, 1972).

65. Shoven, J.B. "General Equilibrium with Taxes: A Computational Procedure and an Existence Proof," Review of Economic Studies, XL (October, 1973).

66. Slutsky, E. "Sulla teoria del bilancio del consumatore," Gior. degli econ., LI, 1915.

67. Smith, V.L. "Economics of Production from Natural Resources," American Economic Review, LVIII (June, 1968).

68. Smith, V.L. "An Optimistic Theory of Exhaustible Resources," Journal of Economic Theory, IX (December, 1974).

69. Solow, R.M. "The Economics of Resources or the Resources of Economics," American Economic Review, Papers and Proceedings, LXIV (May, 1974).

70. Solow, R.M. "Intergenerational Equity and Exhaustible Resources," Review of Economic Studies Symposium, (summer, 1974).

71. Spiegal, M.R. Advanced Calculus, New York: Shaum, 1963.

72. Stigler, G.J. and K. E. Boulding. Readings in Price Theory, Chicago: Irwin, 1952.

73. Surrey, A.J. and W. Page. "Some Issues in the Current Debate about Energy and Natural Resources," in Pearce and Rose, Economics of Natural Resource Depletion, London: Macmillan, 1975.

74. Tullock, G. "The Social Rate of Discount and the Optimal Rate of Investment: Comment," Quarterly Journal of Economics, LXXVIII (May, 1964).

75. Van Long, Ngo. "The Theory of the Mineral Firm and the Optimal Extraction of a Non-Renewable Resource," Australian Economic Papers, XIII (December, 1974).

76. Vickery, W. "The Economic Criteria for Optimum Rates of Depletion," in Mason Gaffney, Extractive Resources and Taxation, Madison, Wisconsin: University of Wisconsin Press, 1967.

77. Vousden, N. "Basic Theoretical Issues of Resource Depletion," Journal of Economic Theory, VIII (December, 1973).

78. Weinstein, M. and R. Zeckhauser. "Use Patterns for Depletable and Recyclable Resources," Review of Economic Studies Symposium, (summer, 1974).

79. Weinstein, M. and R. Zeckhauser. "Optimal Consumption of Depletable Resources," Quarterly Journal of Economics, LXXXIX (August, 1975).

80. Willig, R.D. "Consumer's Surplus Without Apology," American Economic Review, LXVI (September, 1976).

APPENDIX A

The Implicit Function Theorem

According to the Implicit Function Theorem[1] a system of n equations containing n endogenous variables and m parameters:

A-1
$$\begin{aligned} &F^1(x_1, \ldots x_n; a_1, \ldots a_m) = 0 \\ &F^2(x_1, \ldots x_n; a_1, \ldots a_m) = 0 \\ &\cdot \\ &\cdot \\ &\cdot \\ &F^n(x_1, \ldots x_n; a_1, \ldots a_m) = 0 \end{aligned}$$

If: a. all the functions $F^1, \ldots F^n$ have continuous partial derivatives with respect to the x_i's and the a_j's and

b. if at a point $(x_{1_0}, \ldots x_{n_0}, \ldots a_{m_0})$ satisfying A-1, the following Jocobian determinant is nonzero:

$$|J| = \begin{vmatrix} \frac{\partial F^1}{\partial x_1} & \ldots & \frac{\partial F^1}{\partial x_n} \\ \vdots & & \\ \frac{\partial F^n}{\partial x} & \ldots & \frac{\partial F^n}{\partial x_n} \end{vmatrix}$$

Then: There exists an m-dimensional neighborhood of $(a_{1_0}, \ldots a_{m_0})$, N, in which the endogenous variables $x_1, \ldots x_n$ are functions of the exogenous variables $a_1, \ldots a_m$:

---

[1]See Alpha C. Chiang, Fundamental Methods of Mathematical Economics, second edition, (McGraw-Hill, 1974), pp. 218-27.

A-2
$$x_1 = f^1(a_1, \ldots a_m)$$
$$\vdots$$
$$x_n = f^n(a_1, \ldots a_m)$$

Moreover, it can be shown that:

A-3
$$\begin{bmatrix} \frac{\partial F^1}{\partial x_1} & \cdots & \frac{\partial F^1}{\partial a_1} \\ \vdots & & \vdots \\ \frac{\partial F^n}{\partial x_1} & \cdots & \frac{\partial F^n}{\partial x_n} \end{bmatrix} \begin{bmatrix} \frac{\partial x_1}{\partial a_1} \\ \\ \\ \frac{\partial x_n}{\partial a_1} \end{bmatrix} = \begin{bmatrix} \frac{\partial F^1}{\partial a_1} \\ \vdots \\ \frac{\partial F^n}{\partial a_1} \end{bmatrix}$$

Typically, in economic problems of comparative statics, we are interested in obtaining the column vector:

$$\frac{\partial x_i}{\partial a_j}, \; i = 1, \ldots n \; ; \; j \text{ given}$$

when we only have the equations given in the form of A-1. From the expression A-3 we can obtain this column vector through matrix inversion, or more simply, by using Cramer's Rule.

APPENDIX B

Given the seven equations, in implicit form:

$$F^i = F^i(R_{1_o}, R_{2_o}, T, T_1, T_2, P_{1_o}, P_{2_o}, m_1, m_2, c_1, c_2, r) = 0;\ i = 1, \ldots 7$$

$$\begin{aligned}
\text{where } F^1 &= R_{1_o} e^{rT_1} + c_1 - (R_{2_o} + c_2) &&= 0\\
F^2 &= f(R_{2_o} e^{rt} + c_2) &&= 0\\
F^3 &= T - T_1 - T_2 &&= 0\\
F^4 &= P_{1_o} - R_{1_o} - c_1 &&= 0\\
F^5 &= P_{2_o} - R_{2_o} - c_2 &&= 0\\
F^6 &= \int_o^{T_1} f(R_{1_o} e^{rt} + c_1)dt - m_1 &&= 0\\
F^7 &= \int_{T_1}^{T} f(R_{2_o} e^{rt} + c_2)dt - m_2 &&= 0
\end{aligned}$$

where $R_{1_o}, R_{2_o}, T, T_1, T_2, P_{1_o}, P_{2_o}$ are seven endogenous variables and m_1, m_2, c_1, c_2, r are five parameters, it is desired to find the value of the column vectors:

$$\left[A_j\right] = \begin{bmatrix} \frac{\partial R_{1_o}}{\partial m_1} \\ \frac{\partial R_{2_o}}{\partial m_1} \\ \frac{\partial T}{\partial m_1} \\ \frac{\partial T_1}{\partial m_1} \\ \frac{\partial T_2}{\partial m_1} \\ \frac{\partial P_{1_o}}{\partial m_1} \\ \frac{\partial P_{2_o}}{\partial m_1} \end{bmatrix}; \begin{bmatrix} \frac{\partial R_{1_o}}{\partial m_2} \\ \frac{\partial R_{2_o}}{\partial m_2} \\ \frac{\partial T}{\partial m_2} \\ \frac{\partial T_1}{\partial m_2} \\ \frac{\partial T_2}{\partial m_2} \\ \frac{\partial P_{1_o}}{\partial m_2} \\ \frac{\partial P_{2_o}}{\partial m_2} \end{bmatrix}; \begin{bmatrix} \frac{\partial R_{1_o}}{\partial c_1} \\ \frac{\partial R_{2_o}}{\partial c_1} \\ \frac{\partial T}{\partial c_1} \\ \frac{\partial T_1}{\partial c_1} \\ \frac{\partial T_2}{\partial c_1} \\ \frac{\partial P_{1_o}}{\partial c_1} \\ \frac{\partial P_{2_o}}{\partial c_1} \end{bmatrix}; \begin{bmatrix} \frac{\partial R_{1_o}}{\partial c_2} \\ \frac{\partial R_{2_o}}{\partial c_2} \\ \frac{\partial T}{\partial c_2} \\ \frac{\partial T_1}{\partial c_2} \\ \frac{\partial T_2}{\partial c_2} \\ \frac{\partial P_{1_o}}{\partial c_2} \\ \frac{\partial P_{2_o}}{\partial c_2} \end{bmatrix}; \begin{bmatrix} \frac{\partial R_{1_o}}{\partial r} \\ \frac{\partial R_{2_o}}{\partial r} \\ \frac{\partial T}{\partial r} \\ \frac{\partial T_1}{\partial r} \\ \frac{\partial T_2}{\partial r} \\ \frac{\partial P_{1_o}}{\partial r} \\ \frac{\partial P_{2_o}}{\partial r} \end{bmatrix}$$

$j = 1, \ldots 5$

According to the Implicit Function Theorem:

$$\begin{bmatrix} A_1 \end{bmatrix} = \begin{bmatrix} \dfrac{\partial R_{1_o}}{\partial m_1} \\ \dfrac{\partial R_{2_o}}{\partial m_1} \\ \dfrac{\partial T}{\partial m_1} \\ \dfrac{\partial T_1}{\partial m_1} \\ \dfrac{\partial T_2}{\partial m_1} \\ \dfrac{\partial P_{1_o}}{\partial m_1} \\ \dfrac{\partial P_{2_o}}{\partial m_1} \end{bmatrix} \begin{bmatrix} \dfrac{\partial F^1}{\partial R_{1_o}} & \cdots & \dfrac{\partial F^1}{\partial P_{2_o}} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \dfrac{\partial F^7}{\partial R_{1_o}} & & \dfrac{\partial F^7}{\partial P_{2_o}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial F^1}{\partial m_1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial F^7}{\partial m_1} \end{bmatrix}$$

A similar expression can be found for the other column vectors,

$$\begin{bmatrix} A_j \end{bmatrix} \quad , j = 2, \ldots 5.$$

Referring to the above matrix equation as $\begin{bmatrix} A_i \end{bmatrix} \begin{bmatrix} B \end{bmatrix} = \begin{bmatrix} C_i \end{bmatrix}$

$$i = 1, \ldots 5$$

the determinant of B,

$$\left| B \right| = \begin{vmatrix} x_1 & -1 & 0 & x_2 & 0 & 0 & 0 \\ 0 & x_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 \\ x_5 & 0 & 0 & x_6 & 0 & 0 & 0 \\ 0 & x_7 & x_9 & x_8 & 0 & 0 & 0 \end{vmatrix} \quad .$$

where the x_i and their algebraic signs are as follows:

$$x_1 = e^{rT_1} \qquad = +$$

$$x_2 = rR_{1_o} e^{rT_1} \qquad = +$$

$$x_3 = f^1(R_{2_o} e^{rT} + c_2)e^{rT} \qquad = -$$

$$x_5 = \int_o^{T_1} f^1(R_{1_o} e^{rt} + c_1)e^{rt} \qquad = -$$

$$x_6 = f(R_{1_o} e^{rT_1} + c_1) \qquad = +$$

$$x_7 = \int_{T_1}^{T} f^1(R_{2_o} e^{rt} + c_2)e^{rt} \qquad = -$$

$$x_8 = -f(R_{2_o} e^{rT_1} + c_2) \qquad = -$$

$$x_9 = f(R_{2_o} e^{rT} + c_2) \qquad = +$$

According to Cramer's Rule, each term in the column vector, A_j , j = 1, . . . 5, can be calculated by dividing a uniquely-modified B-determinant by the determinant of B itself. In order to calculate the ith term of A_j , j = 1, . . . 5, simply replace the its column of $|B|$ with the column vector C_j , j = 1, . . . 5;

Proceeding in this fashion, the entire vector, A_j , $j = 1, . . . 5$, can be determined.

From the algebraic signs of the terms in $|B|$, the algebraic sign of the terms in the five A_j can be shown to equal the following:

$$\left[A_1 \right] = \begin{bmatrix} \frac{\partial R_{1_o}}{\partial m_1} \\ \frac{\partial R_{2_o}}{\partial m_1} \\ \frac{\partial T}{\partial m_2} \\ \frac{\partial T_1}{\partial m_1} \\ \frac{\partial T_2}{\partial m_1} \\ \frac{\partial P_{1_o}}{\partial m_1} \\ \frac{\partial P_{2_o}}{\partial m_1} \end{bmatrix} = \begin{bmatrix} - \\ 0 \\ + \\ + \\ 0 \\ - \\ 0 \end{bmatrix} \; ; \; \left[A_2 \right] = \begin{bmatrix} \frac{\partial R_{1_o}}{\partial m_2} \\ \frac{\partial R_{2_o}}{\partial m_2} \\ \frac{\partial T}{\partial m_2} \\ \frac{\partial T_1}{\partial m_2} \\ \frac{\partial T_2}{\partial m_2} \\ \frac{\partial P_{1_o}}{\partial m_2} \\ \frac{\partial P_{2_o}}{\partial m_1} \end{bmatrix} = \begin{bmatrix} - \\ - \\ + \\ + \\ + \\ - \\ - \end{bmatrix}$$

$$\left[A_3\right] = \begin{bmatrix} \dfrac{\partial R_{1_o}}{\partial c_1} \\[2ex] \dfrac{\partial R_{2_o}}{\partial c_1} \\[2ex] \dfrac{\partial T}{\partial c_1} \\[2ex] \dfrac{\partial T_1}{\partial c_1} \\[2ex] \dfrac{\partial T_2}{\partial c_1} \\[2ex] \dfrac{\partial P_{1_o}}{\partial c_2} \\[2ex] \dfrac{\partial P_{2_o}}{\partial c_2} \end{bmatrix} = \begin{bmatrix} - \\[2ex] 0 \\[2ex] + \\[2ex] + \\[2ex] 0 \\[2ex] + \\[2ex] 0 \end{bmatrix} ; \left[A_4\right] = \begin{bmatrix} \dfrac{\partial R_{1_o}}{\partial c_2} \\[2ex] \dfrac{\partial R_{2_o}}{\partial c_2} \\[2ex] \dfrac{\partial T}{\partial c_2} \\[2ex] \dfrac{\partial T_1}{\partial c_2} \\[2ex] \dfrac{\partial T_2}{\partial c_2} \\[2ex] \dfrac{\partial P_{1_o}}{\partial c_2} \\[2ex] \dfrac{\partial P_{2_o}}{\partial c_2} \end{bmatrix} = \begin{bmatrix} + \\[2ex] - \\[2ex] + \\[2ex] + \\[2ex] + \\[2ex] + \\[2ex] + \end{bmatrix} ; \left[A_5\right] = \begin{bmatrix} \dfrac{\partial R_{1_o}}{\partial r} \\[2ex] \dfrac{\partial R_{2_o}}{\partial r} \\[2ex] \dfrac{\partial T}{\partial r} \\[2ex] \dfrac{\partial T_1}{\partial r} \\[2ex] \dfrac{\partial T_2}{\partial r} \\[2ex] \dfrac{\partial P_{1_o}}{\partial r} \\[2ex] \dfrac{\partial P_{2_o}}{\partial r} \end{bmatrix} = \begin{bmatrix} - \\[2ex] - \\[2ex] - \\[2ex] - \\[2ex] - \\[2ex] - \\[2ex] - \end{bmatrix}$$

These values are presented in a somewhat different form in the body of the dissertation. See model II.

APPENDIX C

Given the four equations, in implicit form:

$$F^1(T,R_o,P_o,U;C_B) = \int_o^T f(R_o e^{rt} + c)dt - m = 0;\ f(.) < 0$$

$$F^2(T,R_o,P_o,U;C_B) = R_o e^{rt} + C - C_B = 0$$

$$F^3(T,R_o,P_o,U;C_B) = P_o - R_o - C = 0$$

$$F^4(T,R_o,P_o,U;C_B) = \int_o^T g(f(R_o e^{rt} + c))e^{-rt}dt - U = 0;\ g(.) > 0$$

where T, R_o, P_o, U are four endogenous variables and C_B is the relevant parameter, it is desired to find the value of the column vector:

$$\begin{bmatrix} A \end{bmatrix} = \begin{bmatrix} \frac{\partial T}{\partial C_B} \\ \frac{\partial R_o}{\partial C_B} \\ \frac{\partial P_o}{\partial C_B} \\ \frac{\partial U}{\partial C_B} \end{bmatrix}$$

According to the Implicit Function Theorem:

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} \frac{\partial F^1}{\partial T} & \cdots & \frac{\partial F^1}{\partial U} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \frac{\partial F^4}{\partial T} & & \frac{\partial F^4}{\partial U} \end{bmatrix} = \begin{bmatrix} -\frac{\partial F^1}{\partial C_B} \\ \cdot \\ \cdot \\ \cdot \\ -\frac{\partial F^4}{\partial C_B} \end{bmatrix}$$

Referring to the matrix equation C_5 as $A \; B \; = \; C$, the determinant of B is equal to:

$$\begin{vmatrix} B \end{vmatrix} = \begin{vmatrix} X_1 & X_2 & 0 & 0 \\ X_3 & X_4 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ X_5 & X_6 & 0 & -1 \end{vmatrix}$$

where

$$X_1 = \int_o^T f^1(R_o e^{rt} + c)e^{rt}$$

$$X_2 = f(R_o e^{rT} + c)$$

$$X_3 = e^{rT}$$

$$X_4 = rR_o e^{rT}$$

$$X_5 = \int_o^T \frac{dg}{df} f^1(R_o e^{rt} + c)$$

$$X_6 = g(f(R_o e^{rT} + c))e^{-rT}$$

The algebraic signs of $X_1, \ldots X_6$ are as follows:

$$X_1 = -$$
$$X_2 = +$$
$$X_3 = +$$
$$X_4 = +$$
$$X_5 = -$$
$$X_6 = +$$

According to Cramer's Rule, each term in the column vector A, can be calculated by dividing a uniquely modified B-determinant by the determinant of B itself. In order to calculate the ith term of A, simply replace the ith column of B with the column vector C. Proceeding in this fashion, the entire vector, A, can be calculated.

Given the algebraig signs of the elements in $|B|$, summarized by $X_1, \ldots X_6$, the algebraic signs of the elements in the A vector are as follows:

$$\begin{bmatrix} \frac{\partial T}{\partial C_B} \\ \frac{\partial R_o}{\partial C_B} \\ \frac{\partial P_o}{\partial C_B} \\ \frac{\partial U}{\partial C_B} \end{bmatrix} = \begin{bmatrix} + \\ + \\ + \\ + \end{bmatrix}$$