Frontiers of usefulness: The economics of exhaustible resources

by R.C.A. Minnitt*

Synopsis

The influence of Harold Hotelling's work in the field of mineral economics is both unusual and outstanding. In the 66 years since its development, theory surrounding the economics of exhaustible resources has only been significantly researched since the early 1970s. At that time a wave of research into the economics of exhaustible resources arose from the threat of global resource scarcity and continued into the late 1990s. Current East Asian and global demand for natural resources as well as mineral and metal commodities, has stimulated interest in the optimal and sustainable exploitation of exhaustible resources to the point that the subject is worthy of review. Hotelling's theory and the assumptions that he made in establishing a model for the optimal exploitation of mineral resources at the industry level and at the level of the individual mine owner, are examined. He showed that the rate of increase in exhaustible resource prices should be equivalent to the market rate of interest and that in a market with rising prices the rate of extraction should be constant. In markets where demand is stable optimal exploitation requires that the rate of extraction should decline over time, while in a monopoly market he showed that the rate of extraction is retarded. In addition, the effects of cumulative production on costs and prices, the uncertainty associated with estimates of the resource stock, the impact of exploration and issues of intergenerational equity are examined. An analysis of the optimal price path, extraction costs and mineral rents at an industry level is presented as a model. An analysis of the rates of production and optimal benefits at the individual mine level including total, average and marginal costs on the optimal rate of production, is considered. Optimal extraction rates for a mine with a clearly defined concave positive cost equation vary according to the position on the marginal cost curve. The influence of discount rates and the size of the ore reserves on the rates of mineral extraction are also examined and could define the frontier of usefulness for the 'economics of exhaustible resources'.

Introduction

Annually the content and relevance of the postgraduate course in Mineral Economics in the School of Mining Engineering at the University of the Witwatersrand is reviewed. The connection between classical economics and exhaustible resource depletion is well established, but at the end of teaching such a course, especially to mining people, the benefits of the classical economic approach seem less convincing. However, the classical insights about rent and the optimal depletion of exhaustible resources, provided by Harold Hotelling and other great resource economists, remain one of the pillars in the course on Mineral Economics.

The reason for the concern is that the extensive literature on the theory of exhaustible resources, is evidence of its interest to academics, but it does not provide much in the way of usable concepts to make mining people better extractors and depleters of natural resources. Postgraduate students working in the minerals industry generally have a technical rather than an economic bent, so why teach economic concepts relating to the extraction of exhaustible resources and where does the frontier of usefulness to the mining industry end? The answer lies in a quotation from Harold Hotelling (1931) that could have appeared in the latest issue of an economics or popular news magazine:

'Contemplation of the world's disappearing supplies of minerals, forests, and other exhaustible assets has led to demands for regulation of their exploitation. The feeling that these products are now too cheap for the good of future generations, that they are being selfishly exploited at too rapid a rate, and that in consequences of their excessive cheapness they are being produced and consumed wastefully has given rise to the conservation movement.'

What's more, the relevance of his words grows with time—what was true in 1931, is much truer in 2007.

Two reasons for continuing to teach the principles that Hotelling unfolded to us are firstly, the balance brought to the economic overemphasis by a little-known article by Paul

* *School of Mining Engineering, University of the Witwatersrand.*

 SA ISSN 0038–223X/3.00 + 0.00. Paper received March 2007. Revised paper received June 2007.

Bradley (1985) who asks the question, 'Has the economics of exhaustible resources helped the economics of mining?'. The answer is skewed to emphasize a totally different set of issues and problems to those raised by Hotelling. The second reason is that with the rise of the concept of sustainable development in the mining and minerals industry, there is value in understanding arguments about socially optimal rates of depletion, discount rates, and the way in which markets are regulated. This is not through public policy as much as by directors of the global minerals industry who in their own best interests are probably doing as good a job of ensuring the long-term sustainability of benefits from natural resources as can be expected.

The aim of this paper is to revisit the work of Harold Hotelling and, having established his main premises, to examine how the principles apply at the minerals industry level where all producers face a stable demand curve and the level of the individual firm or mine, and to glean from them concepts that can be beneficially applied to mineral developments. Apart from the different views Bradley (1985) raises about issues that are critical to the mining industry, the value of Hotelling's work has remained unquestioned by most researchers in mineral economics.

The work of Harold Hotelling

The work of Hotelling written in 1931 remained largely unexplored for about 40 years because of his style and mathematical approach to the issues. Only in 1973, when his work was popularized by Robert Solow in his Richard T. Ely Lecture entitled the 'Economics of Resources or the Resources of Economics?', did the writing of Hotelling begin to affect the mineral resources industry. Solow heightened the influence of his work because, being sensitive to the awakening concerns about resource scarcity, he presented it just after the OPEC announced the increased in oil prices in 1972. The next decade was to produce an unprecedented flow of literature dealing with issues about the allocation and depletion of resources.

Hotelling's model predicts a general rise in commodity prices over time and even though we have not observed such increases in prices until recently (the last five years), the model has been used by numerous authors as a useful reference point in discussions on the various dimensions of mineral supply and availability (Eggert, 2007). Among the points that the model helps introduce are that:

- Prices are a useful indicator of scarcity, if markets are functioning well
- The effects of exploration and technological innovation significantly and importantly influence mineral availability over time
- Market structure matters (competition versus monopoly)
- Mineral resources are not homogeneous
- Backstop technologies limit the degree to which prices can increase
- Substitution is an important response to increased scarcity
- Changes in demand influence price and availability.

In other words, the model provides a vehicle for introducing the various dimensions of mineral supply and scarcity (Eggert, 2007).

In an article entitled 'Hotelling's "Economics of Exhaustible Resources": Fifty year later', Devarajan and Fisher (1981) celebrate the importance of Hotelling's article on its fiftieth anniversary. Their aim was to show how Hotelling (1931), in an article that had been neglected for so long, had answered many questions that arose about a concern for the adequacy of the world's natural resources in the decade between 1972 and 1981. They point out that Hotelling's reasons for writing the paper were firstly to assess the policy debates arising out of the conservation movement and secondly to develop an adequate theory for the exploitation of exhaustible natural resources.

Hotelling's assumptions

Hotelling (1931) established several economic assumptions for his extractive model for natural resources that simply do not reflect truth in the real world of minerals extraction. Considering the assumptions on which his rule is based, it is hardly surprising that the rule cannot be verified empirically. The model assumes firstly, that the mine owner's objective is to maximize the present value of its current and future profits. This requires that extraction take place along an efficient path in a competitive industry equilibrium, which implies that all mines are identical in terms of costs and that they are all price takers in a perfect and instantaneous market of information. Secondly, the mine is perfectly competitive and has no control over the price it receives for its production. A third assumption is that mine production is not constrained by existing capacity; it may produce as much or as little as it likes at any time during the life of the mine. The forth assumption is that the ore deposit has a capitalized value. That is, a copper or gold deposit in the ground is a capital asset to its owner (and society) in the same way as any other production facility. Furthermore, he assumed that the richest and most accessible deposits would be mined first, and that increasing scarcity (after exhaustion of the best mines) would confer capitalized value on inferior deposits, which could then be mined. Fifthly, he assumed the resource stock is homogenous and consequently there is no uncertainty about the size, grade and tonnage of the ore deposit. Current and future prices and extraction costs are known. This implies that an orebody has uniform quality or grade throughout and that there is no change in grade of the ore as mining proceeds. Miners and grade control officers, who endeavour to supply the mill only with ore above a certain grade, recognize this fifth assumption to be major departure from reality. The sixth assumption is that the costs of mining or extraction do not change as the orebody is depleted. Again this assumption does not recognize that all mines face increasing costs as the ores are depleted. Underground mining costs increase as the mining face becomes longer and deeper and moves further away from the shaft system, while in open pit operations haul roads become longer and pits become progressively larger and deeper. A rider to Hotelling's assumption that the marginal unit

(standard mining unit, smu) is accessible at the same constant cost, is the assumption that the marginal cost of extraction in this particular case is zero. In addition, it implies that the market price and the rate of extraction are connected by a stable, downward sloping demand curve for the resource (Solow, 1974, p. 3). In this constrained model the size of the remaining stock declines without ever being augmented by exploration discoveries. The final assumption is that there is no technological improvement during the life of the mine and that no new additions to the resource stock are contributed by exploration. While the concepts that Hotelling applied are enticing to the economist, his assumptions tend to diminish the potential value of the application for the miner involved in mineral developments.

However, the volume of literature that has proliferated on the diverse aspects of the economics of exhaustible resources, suggests it is an academically appealing theory. Herein lies the problem. It is principally a theoretical economic construct that provides huge insight into a realm of modelled economics where everything behaves in an 'economic' manner. In their conclusion Devarajan and Fisher (1981) state that Hotelling's 1931 article is 'the sole source of work in a vigorously growing branch of economics.' (p.71). His elegant analysis, asides, conjectures and canonical model provide economists with a structure to build on that is almost a generation ahead of its time. Solow (1974) stated that, 'Good theory is usually trying to tell you something, even if it is not the literal truth.' (p. 10), so although the economics of exhaustible resources does not invade the real world of mining and mineral extraction to any large extent, it is still worthwhile to re-examine the theory.

A wave of research output arose from Hotelling's work as many investigators relaxed his early assumptions and introduced a flexibility that widened the scope of the model applications. However, as with most highly constrained economic models, relaxing the assumptions magnifies the uncertainty associated with the outputs of the model. Hotelling's approach to the depletion of exhaustible resources can be applied at both the industry and the individual mine level, and in this paper application of Hotelling's work at both levels is considered.

Optimal extraction of exhaustible resources at industry level

The price of an exhaustible resource must grow at the rate of interest

Solow (1972) pointed out that a mineral deposit, whose value arises from the potential for extraction and sale, is a capital asset to its owner and society. The main difference is that natural resources are not reproducible and that the size of the existing stock can never increase, only decease through time. Even with the possibility of some recycling, metal commodities would remain an exhaustible resource.

According to Solow the importance of Hotelling's work is based on his assertion that 'the only way a resource in the ground can produce a current return for its owner is by appreciating in value' (p. 2). This has become the famous 'Hotelling r-per cent rule' which specifically states that the price of an exhaustible resource (or value of an exhaustible resource in the ground) must grow at the market rate of interest, and is given by the equality:

$$p_t = p_0 e^{rt}$$

where p_t is the price in period t, p_O is the price in the initial period, and r is the market rate of interest. This of course holds only if extraction costs are zero. More generally, net price rises at the rate of interest for all versions of the model (zero extraction costs, constant positive costs, rising costs). It is true even under monopoly, in which case marginal revenue less marginal extraction cost rises at the rate of interest. The value of a mineral deposit is also the present value of future sales from it, less the costs of extraction. Thus the net price (λ), which is the market price minus the costs of extracting one ton of ore (marginal cost), i.e. $\lambda = p - mc$, should be expected to grow exponentially at the rate of interest. The marginal costs should include the amount paid in fixed and variable costs, wages, taxes, dividends, and fair market return, sufficient to induce the mine owner to invest in a mineral development in the first place. This simple case of net price has been extended to mean that the royalty (the price net of the extraction costs) for the marginal unit of extraction, will rise at the rate of interest. This is exactly the definition of royalty. It is the decline in value of a natural resource as a result of the extraction of one unit of the resource, the payment due to the owner for the exploitation of his resource. The miner should in this case be happy to pay the royalty to the owner of the resource, because it is a benefit he never aimed for. In reality mine owners are loath to pay this amount to the owner of the resource, usually in the form of a royalty to the government, as they view it as a return to be appropriated because of their entrepreneurial skill and the risk they bear in mineral investment ventures.

The outcome of appreciating resource value was also achieved by Gray (1914) who assumed that resource prices remained constant over time, while the marginal costs of extraction rise. The approach was later investigated by Carlisle (1954) who concluded that the growing difference between mineral price and the costs of extraction was exactly equal to the market rate of interest (Devarajan and Fisher, 1981, p.66). If as the theory states, prices grow at the rate of interest, they cannot grow indefinitely. There must come a time when prices are so high that production falls to zero and, according to the theory, the last ton produced will also be the last ton in the ground. As Solow (1974) states, 'The resource will be exhausted the instant it has priced itself out of the market' (p. 3). It should however be noted that there is a considerable body of research including work by Arrow and Chang (1978), that sets out to examine the r-per cent rule, simply because prices do not behave as Hotelling predicted !

If there are two sources for the same mineral production, it is impossible for them both to operate under optimal conditions of output, unless their costs of extraction are identical. If their costs of extraction differ, first the low cost and then the high costs producer will operate. In this sense there is not just one market, but a sequence of markets with

successively higher prices, from now until the time of exhaustion (Solow, 1974). The extension of this idea is that the lowest cost producer is the first to operate then, as market prices rise and Resource 1 is depleted, the next lowest cost mine comes into production, and so on until all the ore, even from the highest cost producer is exhausted, at which time production of that particular mineral will cease. Although this sequence would maximize the welfare and benefits to society and the mine owners, there is no way of firstly ordering the deposits according to extraction costs, even for a given region this way, and secondly, of persuading owners to deplete their ores only in a given sequence.

The concept that a 'backstop technology' (Nordhaus, 1973) or high cost substitute to replace production from the natural resource, for example nuclear power replacing oil as an energy source, is invoked once the resource is totally exhausted and market prices have risen to a point where the cost of the technology becomes competitive. This concept is particularly relevant in the current global energy market where uranium has risen from less than US$5 per pound to US$115 per pound as the prospect of diminished oil supplies and growing insecurity of supply looms larger than ever. Thus, the cost of the backstop technology provides a ceiling price for the natural resource (Solow, 1974, p. 4).

Solow (1974) and Hartwick and Olewiler (1998) also differentiate between resources that occur in stock and those that occur as mineral flows. Generally the flow markets are considered unstable, while the markets for mineral stocks provided a corrective force to the markets. The flow equilibrium condition is that net price grows at the market rate of interest. Instability in the flow market is initiated by expectations about future prices, which leads to speculative withholding of supply when higher prices are anticipated and excessive dumping of supply when prices are expected to fall in the longer term. This analysis of spot markets does not take the asset market into account, in which resource owners believe the value of their product is anchored somewhere in the future. If in the longer term prices are not rising at the expected rate, capital losses on assets will be incurred as the value of the stock will have to be written down, or vice versa.

If prices rise the rate of extraction is constant

The 'Hotelling rule' extends to aspects other than just price. The royalty, the undiscounted value of a unit of resource, the scarcity rent, the opportunity cost, the marginal profit minus marginal revenue less marginal costs, and the price net of extraction costs, for the exhaustible resource are all synonymous, and this naturally has implications for the rate of extraction. If the net price (net of extraction costs) is rising at the rate of interest then the discounted value of net price must be the same in all periods, and the undiscounted value of the resource must also grow at the rate of interest. This is usually referred to as the flow condition (Hartwick and Olewiler, 1998). Furthermore, if the net price or value of the resource grows at the rate of interest, there can be no benefit from increasing or decreasing the rate of extraction in any period. In this case demand is exactly equal to supply at every level of production and the market clears at every instant in time (Solow, 1974).

If demand is stable the rate of extraction declines over time to zero

Hotelling's model connects the market price and the rate of extraction by a demand curve for the resource. If demand is stable the rate of extraction must decline monotonically to zero, to the point where the demand curve intersects the price curve at a finite price, as shown in Figure 1. For a competitive firm manufacturing a reproducible good, the rate of production is chosen such that price = *MC* in order to maximize profit. For the owner of a non-renewable resource the equivalent condition is price = *MC* + the opportunity cost of depletion, implying that less of the resource will be extracted in any period, than if it were renewable. This principle lies at the heart of Hotelling's approach and is illustrated in the example presented in Appendix 1 with the results being shown in Table A6 and graphically in Figure 2.

The additional marginal value that scarcity creates is the marginal user cost. The existence of the marginal user cost implies that efficient extraction of ore from a mine will exceed the marginal cost of extraction, creating a scarcity rent for the resources, which is appropriated by the owner of the resource and becomes part of his producer surplus provided property rights are correctly defined (Hartwick and Olewiler, 1986). The introduction of the concept of scarcity affects our understanding of optimal allocations of the resources in highly specific ways. An important aspect of the treatment is that it has to be dynamic, incorporating the impact of time on each of the issues (Lasserre, 1991, p. 1–4).

Scarcity imposes an opportunity cost that must be accounted for in determining how best to allocate resources over time. Opportunity cost of any decision is the foregone value of the next best alternative that was chosen. Choosing mineral development as an investment means that the money is not available for other uses. Furthermore, additional returns above those from mining, which may have accrued from opening say a shoe factory, will be lost.

The additional cost is variously referred to as marginal user cost (from the opportunity cost to the user of mining a tonne of ore today), scarcity rent, royalty, net price, marginal profit, economic rent, shadow price or Hotelling rent. The mine owner must cover the opportunity cost of depletion because a day will come when the ore reserve is completely

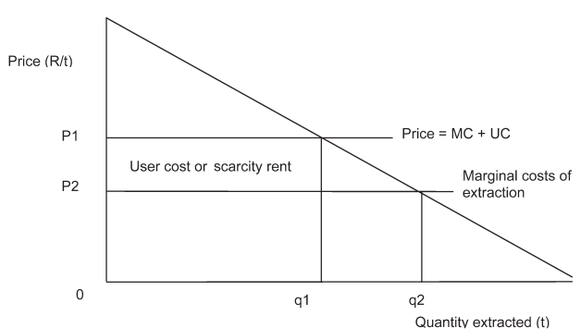

Figure 1—Scarcity rent or marginal user cost for a non-renewable resource

used up. Scarcity implies that the correct and socially desirable price should be higher than the marginal cost of extraction, by an amount equivalent to the opportunity cost, often referred to as the scarcity rent (Figure 1). The fact that scarcity rent is added to the cost means that quantity q_1 is extracted rather than quantity q_2 if there was no scarcity.

The distinction between user costs, which are included as part of the producer surplus and other costs, such as extraction costs, lies in whether or not they are actually paid (Tilton, 2003). Marginal costs of extraction are actually paid, it consumes resources, while the marginal user cost by contrast, is an opportunity cost that would be paid in the form of reduced net benefits if extraction of the resource were not efficiently allocated over time (Lasserre, 1991, p. 1–4). When the allocation is such that net benefits are maximized, this cost is not actually borne, and therefore, the owner of the resource appropriates it as scarcity rent. This source of producer surplus is not eliminated by competition. Consequently, even when time is an important consideration, in the presence of well-defined property rights, market allocations and efficient allocations are identical (Tietenberg, 1988, p.42–45).

Rates of extraction in monopolistic markets are retarded

Devarajan and Fisher (1981) point to the recurrent monopoly episodes in the resource market as good reason for considering the effects of monopoly price on rates of production. It should be pointed out that in reality attempts to monopolize or cartelize the metals markets have proved fruitless and, more often than not, very costly to those who try to maintain them. For some reason the nature and character of the tin market has lent itself to continuous and ongoing attempts at cartelization since 1925 and continued until the mid-1970s to mid 80s when the cartel completely collapsed (Crowson, 2003). Silver was also the subject of cartelization attempt prior to and leading up to 1980 when the Hunt brothers tried to collar the silver market, but again the nature and the unseen movements of the metal behind the open trade market proved to be too unpredictable (Liu, 2002).

Nevertheless Hotelling (1931), held the view that over time a monopoly price would be flatter and rise less rapidly than prices in a competitive market, and that rates of depletion would be slower in monopoly markets than competitive markets, as shown in Figure 2 (Devarajan and Fisher, 1981). The monopolist, who raises price by restricting output, tends to slow the production rate as he depletes his resource, because he has no idea of when the resources might suddenly be exhausted. The competitive resource owner, because he is a price taker, wanting to take advantage of the market price, will try to deplete his resource as fast as possible. Solow (1974) notes the counterintuitive, but amusing sidelight that in such a case the monopolist is in fact the conservationist's friend, something both would be surprised to know.

According to Devarajan and Fisher (1981), the model presented by Hotelling is relaxed to include the possibility of growing potential for substitution in the monopoly market. In early periods the monopolist retards output and takes advantage of the inelastic demand to raise prices. However, if the monopolist suspects that demand for his product may become more elastic (fall), he is likely to accelerate depletion, while the opposite is true if the elasticity of demand increases. As a general impression and in line with Hotelling's conclusions, Devarajan and Fisher (1981) suggest 'monopoly slows depletion' (p. 68). It should be noted that to the extent that monopolists will prefer to slow down production (produce less today than under perfect competition), their interests are aligned with those of the conservationists who are concerned that we are depleting our resources too quickly.

The effects of cumulative production on costs and price

The effects of past production on extraction costs and on demand price have been examined by a number of other researchers, among them Herfindahl (1967), Heal (1976), Solow and Wan (1976), Weitzman (1976), and Hartwick (1978). In simple terms, extraction costs rise as the cumulative stock grows and provided the stock depreciates in

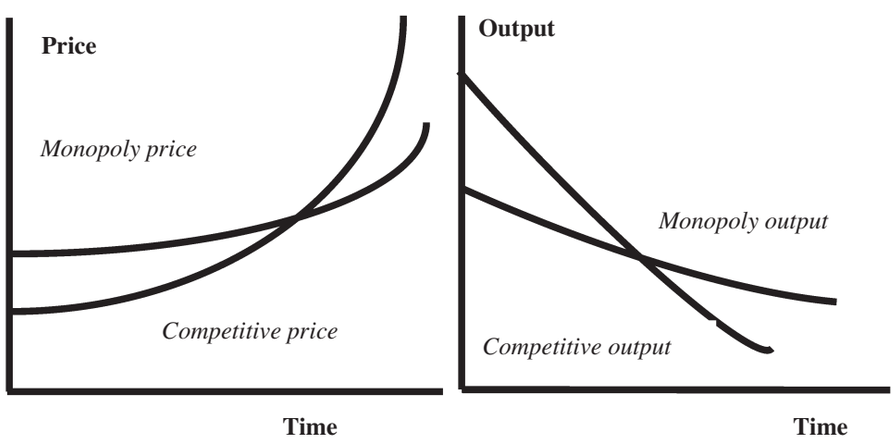

Figure 2—Price and output trajectories for competitive and monopoly markets (after Devarajan and Fisher, 1981, p. 67)

value through some form of 'rusting', Hotelling's r-per cent rule still holds. If the stock does not depreciate the price will be forced down instead of rising at the rate of interest.

In the combined case of rising extraction costs and falling demand, price will fall, but where the stock depreciates the price will follow a U-shaped trajectory (Levhari and Pindyck, 1979). Devarajan and Fisher comment on the research undertaken by Pindyck (1978), who showed that if the stock is augmented by exploration discoveries, price falls as the major finds are made, but eventually it will rise.

The effects on costs of production as a stock of cumulative production grows was first noted by David Ricardo (1817) who in his analysis of mines recognized that like agricultural land, mineral resources would occur in large supply, but at quite different degrees of grade. Ricardo noted that as richer orebodies are depleted, lower grade deposits with higher extraction costs would be exploited. Hotelling noted that the profit that the owner of a mine producing durables such as gold or diamonds can make will depend on his current rate of production, as well as the accumulated stock in circulation. The reasons for this, according to Devarajan and Fisher (p. 69), is that extraction costs increase as the mine becomes deeper and that the demand for durables such as gold and diamonds is affected by the amount of accumulated stock in circulation.

The effects of uncertainty about resource estimates

Hotelling examined only one question about the uncertainty of resource estimates, asking: 'What is the value of a mine, when its contents are supposedly fully known and, what is the effect of uncertainty of estimate?' (p. 139). Fifty years after Hotelling's work, Devarajan and Fisher (1981) are in a far better position to analyse the rich variety of questions that has arisen as a consequence of developments in decision-making under uncertainty. They indicate that Hotelling never answered this question, but instead examined exploration uncertainty and its implications for public policy (p.70). He mentions the great wastes that arise from sudden and unexpected mineral discoveries leading to wild rushes, immensely wasteful socially, to claim valuable property. It appears from the statements by Devarajan and Fisher that Hotelling was concerned about the ability of a discoverer of a mineral deposit to file a claim and so exclude competitors from access. This in turn leads to socially excessive (and therefore wasteful) levels of exploration activity.

Hotelling also recognized that owners on adjoining land could benefit from a discovery following exploration carried out on a neighbour's property, without ever investing in exploration activity themselves. This so-called spillover of geological knowledge leads to higher land prices and excess profits, simply by observing the outcome of his neighbour's exploration activity. Hotelling believes that government as custodians of the national patrimony have a right to expropriate excess profits from the owner on behalf of the nation and not allow such windfalls to remain in private hands. Such information spillover effects can also lead to socially inefficient levels of exploration because 'everyone waits around hoping his neighbour will drill first' (Devarajan and Fisher, 1981, p. 70).

Uncertainty about the supply and demand for resources has to some extent answered Hotelling's question about the effect of uncertainty about the resource estimate of a mine. Assuming the mine is of uniform grade, the owner extracts ore at a rate that is related to his knowledge about the size of the remaining resource. If he is uncertain about the size of his stock he will tend to deplete slower—the conservative approach—than if he is certain about the estimate, simply because he has no way of knowing when he will run out of ore, much like the motorist driving without a petrol gauge (Devarajan and Fisher, 1981).

Uncertainty can of course be reduced by exploration, or new information about a second stock may become available while depleting the current ore reserves (Hoel, 1978). Several researchers have modelled exploration and likened it to a Poisson process (Arrow and Chang, 1978), a continuous stochastic process (Pindyck, 1979) and a random production function (Devarajan and Fisher, 1980).

Uncertainty about demand for an exhaustible resource is not conclusive (Devarajan and Fisher, 1981). According to Weinstein and Zeckhauser (1975), a mine owner who is risk-averse will deplete more quickly, hoping to make more money, if prices in the distant future are more uncertain than prices in the near future. If the risk-averse owner believes that he stands to lose as much now as anytime in the future, Lewis (1977) has shown that he will delay extraction because production should be lower in future and therefore the amount at risk will also be lower.

There are three cases in which demand may suddenly disappear, namely, through some technological advance, the introduction of a substitute, or through expropriation of the resource (Devarajan and Fisher, 1981). In such cases Dasgupta and Heal (1974) and Long (1975) have demonstrated that the tendency is to accelerate current production.

Exploration

Mineral discoveries because of exploration, both in the individual mine and at an industry scale, lead to increases in the size of the resource stock. If the rate of discovery exceeds the rate of depletion, stocks can rise despite extraction. Exploration has three basic functions, first to reduce current and future extraction costs. Higher reserves mean lower extraction costs, which would explain why companies explore and hold reserves (Pindyck, 1978). Cairns (1991) is of the opinion that new discoveries are often more costly to extract. Secondly, to generate discoveries so that the mine can continue to extract at a rate compatible with demand. The scarcity rent rises as reserves diminish, and falls when new discoveries are made. Ultimate scarcity depends on the availability of exploration prospects. If there are numerous exploration prospects, resource prices are unlikely to rise on average in the long run. If there are only few prospective exploration targets, resource prices may rise on average, but at a lower rate than predicted by the Hotelling rule, namely the interest rate.

The third function of exploration is that it provides new information about both, (i) undiscovered resources and (ii) extensions (or the lack of) of existing orebodies. Lassere

(1995) showed that the higher the amount of resource discovered (at any date of exploration), the lower the resource price after discovery, reflecting the normal reaction to a lucky discovery. However, the threshold reserve, at which exploration will resume after the discovery, will be higher, reflecting more optimism about future prospects. Such revisions to the stock means that resource prices may diminish with time.

The effects of optimal depletion on intergenerational equity

Although Hotelling does not discuss intergenerational equity per se, much comment about this topic has arisen. Such comment has been incorporated in the thrust for sustainable development with a large and growing field of relevance, especially in the area of minerals development and depletion and from interest groups raising questions about social interest in the pace of exploitation of the world mineral resource endowments. Although the prognosis in 1972 was an irreversible collapse due to natural resource shortages (Meadows and Meadows, 1972), and commodity prices are now sending scarcity signals, we still continue to exploit the earth's natural resources at ever increasing rates.

The main issue at stake is the rate at which mine owners and society would want to discount future consumer surpluses. Only when mines and society choose the same discount rate can 'competitive equilibrium maximise the sum of discounted consumer plus producer surpluses from natural resources' (Solow, 1974, p. 7). According to Hotelling, the optimal solution is skewed if: several owners pump the same oil or gas field, if uncertainty about exploration outcomes leads to wasteful rushes and useless profits, and if large monopolies or oligopolies exist within the extractive industries. If private discount rates are higher than social rates, the tilt of the production equilibrium will be such that scarcity rents and market prices will rise too fast and consequently resources will be exploited too fast and depleted too soon (Solow, 1974).

The two main reasons that private discount rates are higher than social discount rates are firstly an accommodation of personal risk because of a lack of appropriate insurance markets. Two examples are considered by Solow: firstly security of tenure is a type of risk specific to the minerals industry for which there is no insurance, and secondly the taxes on income from capital; 'individuals care about the after tax returns on capital and society about the before tax returns' (Solow, 1974, p. 8).

Example of optimal depletion at industry level

The main conclusions about the optimal exploitation of exhaustible resources that Hotelling arrived at are illustrated using the five period optimization profile shown in Appendix 1 for the depletion of 2500 t of ore over five consecutive periods. The marginal cost of extraction is R200/t, the discount rate is 5 per cent and the linear demand curve for the commodity is $p = 700 - 0.25$ q. The results of an optimal depletion analysis of a non-renewable resource over five periods are shown in Table I and plotted in Figure 3; the following points are relevant:

- The quantity extracted in each period falls over time (Figure 3 and Table I) This implies that the production rate over the life of the mine starts at a maximum and declines over time to zero. In the industry each mine extracts less in each period and because the demand curve is stable, prices rise. Furthermore, all firms are identical and all extraction costs are equal, allowing the extraction path $q_0, q_1, \ldots q_5$ shown in the five period example (Table I, Figure 3, and Appendix 1) to create competitive industry equilibrium in which social welfare (W) is also optimized.
  This is so because W is a maximum when marginal net surplus (mineral rent) is equal across periods, which is Hotelling's r-per cent rule.
- The undiscounted value of mineral rents rises at the market rate of interest over time and by axiom the present value of mineral rents is constant over time, as shown for the five period optimization results in Figure 3. Therefore each firm makes the same profit in terms of present value and each firm is indifferent to the time of extraction. While this may be a condition of Hotelling optimization, the flow of mineral rent is a function of price, an unusually volatile determinant of surpluses. Other so-called shocks that may affect the markets include new information about the volume of reserves, viable competition from substitutable materials, falling costs of competing technologies, near-term political events and drastic movements in price or production (Solow, 1974, p. 7).
- Mineral prices rise at less than the rate of interest. While the 5 per cent rate of interest applies to the rate of increase of the mineral rent the mineral prices only appreciate only at a rate of 3 per cent (see Table A6 in Appendix 1).
- What should be noted is that both the interest rate and the demand curve are taken as given to the mining industry by the rest of the economy. The linear demand curve relates market price of mineral to quantity demands has a negative slope. In out particular example, the demand function $p = 700 - 0.25\,q$, has a choke price of R700 when demand goes to zero, at which point the stock has also been entirely depleted.

Table I

Results of a five period analysis for optimal resource extraction

Period	Tons reserves	Remaining	Rent λ	Price	PV of Rent	Δ λ	Δ p
1	642.74	1857.26	339.32	539.32	323.16	0.05	0.03
2	574.85	1282.41	356.29	556.28	323.16	0.05	0.03
3	503.57	787.84	374.11	574.14	323.16	0.05	0.03
4	428.69	350.15	392.83	592.82	323.16	0.05	0.03
5	350.15	0	412.46	612.45	323.16	0.05	0.03
Total	**2500.0 t**						

Source: Appendix 1

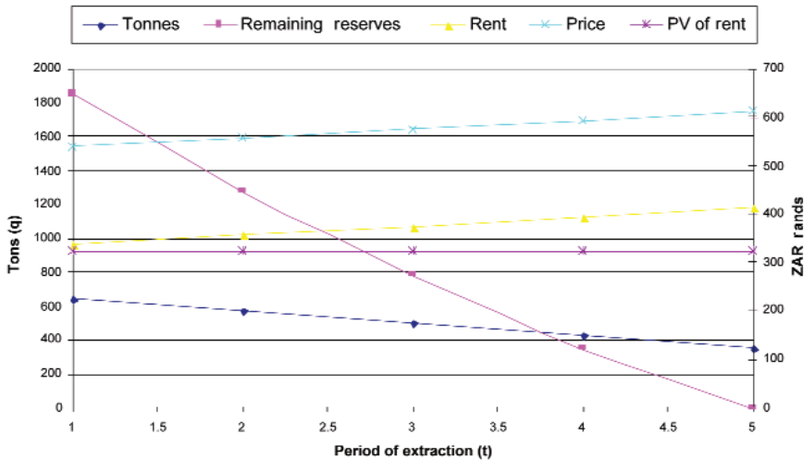

Figure 3—Changes in production rate, remaining reserves, mineral rent (λ), mineral price and the present value of rent for a 2500 t resource under conditions of optimal resource extraction

Optimal extraction at individual mine level

The application of Hotelling's insights at the individual mine level has been explored by a number of investigators, including Hartwick and Olewiler (1998). At this level the Hotelling rule takes a slightly different form in that, rather than the price growing at the rate of interest, we make the assumption that the price remains constant, but the production rate steadily declines over time. The maximization criterion is the same for the industry and individual mines, namely, that the discounted value of the mineral rent, also referred to as the scarcity rent, must remain constant over the life of mine.

Marginal analysis of mines

The principle that competitive firms maximize profits when the rate of production is chosen so that price = *MC*, but profits from extraction of exhaustible resources are maximized when price (*p*) = marginal cost + opportunity cost (rent), is changed in three fundamental ways because the resource is finite.

First, mines are depleting a capital asset, depreciation of which involves a cost, which it is right to account for since there is a lost future opportunity. Thus maximization requires that marginal profit (or the Hotelling rent or scarcity rent), be the same in each period.

Second, the rate of depletion of an exhaustible resource is a type of investment problem. The choice of which asset to invest in, depends on the investors expectation of the rate of return, i.e. an increase in value over time. The marginal return to the mine is the resource rent, namely the value of the ore in the ground. When interest rates are positive, the rent rises at the rate of interest as depletion occurs. If resource rent did not increase in value over time, no one would invest in a mine, because investment in alternative assets would be more valuable. An owner of an existing mine would want to extract all his ore as quickly as technology would allow him, because there would be no point in holding an asset that was not earning at least as much as a saving account could give. Alternatively, if the value of the ore in the ground were appreciating faster than the rate of interest, there would be no incentive to mine. Ore left in the ground would be worth more than an asset (the orebody) reduced in size through mining.

Third, the maximization condition is affected by the stock constraint. This simply means that the amount of ore mined cannot exceed the finite ore reserve. In order to make mining worthwhile, price must exceed the marginal cost by an amount equal to the scarcity rent (Hotelling rent), $p = MC\ (q)$ + scarcity (Hotelling) rent so scarcity rent = $p - MC\ (q)$, must be the same from one period to the next. Furthermore, we can write:

$$\left(\frac{1}{1+r}\right)^{t}\left[p - MC(q_t)\right] = \left(\frac{1}{1+r}\right)^{t+1}\left[p - MC(q_{t+1})\right]$$

or by rearranging

$$\frac{(p - MCq_{t+1}) - (p - MCq_t)}{(p - MCq_t)} = r$$

This is Hotelling's r-per cent rule of extraction.

In order to satisfy the maximization principle, we started off by equalizing marginal profit across any two periods, and ended by showing that $p–MC(q_t)$ increases at a rate equal to *r* per cent. This is illustrated in Figure 4.

Furthermore because $p–MC(q_t)$ is equal to rent, we can say that in order to maximize profits in every period the rent must increase at the rate of interest, *r*. Alternatively we could say the rent in every period must be equal. We can also answer the question, 'What should the rate of production be in order to maximize its profits?' At a rate equivalent to $(q_{(t+1)}–q_t)/q_t$, so that $p–MCq_t$ is increasing at *r* per cent. Rent in this case may be referred to as user cost, royalty, scarcity rent or Hotelling rent. The symbol used for rent is λ, so:

$$p = MC(q_t) + \lambda \text{ or } \lambda = p - MC(q_t)$$

We said earlier that the cost function depends on:

- The rate of production (*q*) and,
- The amount of ore remaining in the ore reserve after each period of extraction (*S*). So $C = f\ (q, S)$.

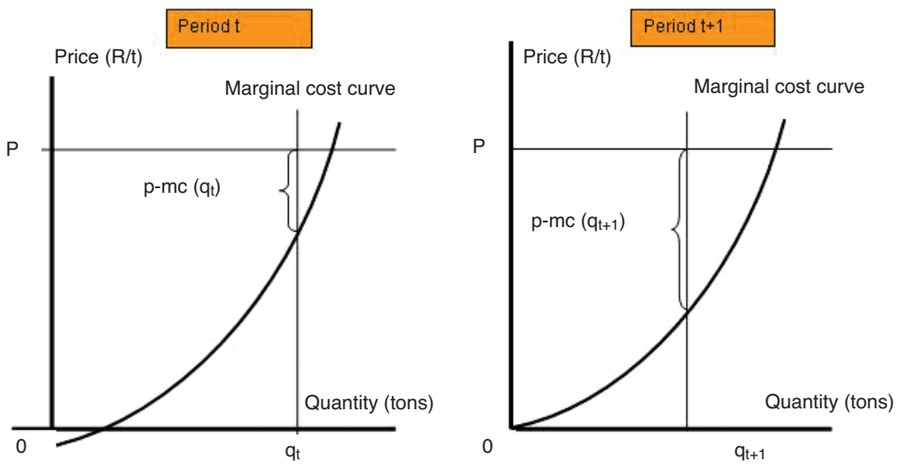

Figure 4—The r per cent rule says that profits are maximized in every period if the rent in each period is the same. Hence rent *p – MC(q_t)* in period *t* must equal rent *p – MC(qt+1)* in period *t*+1 (Modified after Tilton, 2003 and Hartwick and Olewiler, 1998)

Table II

Optimal depletion of the contents of an individual mine

Time years	Rate of production	Remaining reserves	Price AC	Rent (p-AC) (p-AC)q	Average costs life of mine	Annual revenue rent	Net profit over	Discounted
0	324.69	2500.00	5000	1550.00	3450.00	503.27	0.00	1550.00
1	318.86	2175.31	5000	1674.00	3326.00	533.77	53.38	1550.00
2	312.43	1856.45	5000	1807.92	3192.08	564.85	112.97	1550.00
3	305.34	1544.02	5000	1952.55	3047.45	596.19	178.86	1550.00
4	297.49	1238.68	5000	2108.76	2891.24	627.33	250.93	1550.00
5	288.77	941.19	5000	2277.46	2722.54	657.65	328.83	1550.00
6	279.04	652.43	5000	2459.66	2540.34	686.33	411.80	1550.00
7	268.13	373.39	5000	2656.43	2343.57	712.26	498.58	1550.00
8	255.82	105.26	5000	2868.94	2131.06	733.93	587.14	1550.00
	Total	**2422.49**						

We have to ask the question, 'Are our exhaustible resources being depleted too rapidly or too slowly?' This leads to the next question, 'What constitutes optimal use of exhaustible resources?' The question must be considered from the socially efficient and competitive market viewpoints. Other complications, which may arise, include monopoly, cartel behaviour, exploration, technology and substitution, and questions of intergenerational equity.

Example of optimal depletion from a mine

The schedule in Table II illustrates the application of Hotelling's principles to the depletion of reserves from an individual mine. In this case the average cost function is the same as that used in Appendix 2. In addition, for the sake of comparison, a tonnage of 2500 t, used for the industry model has also been applied in this individual mine model, but the reserve is depleted over eight periods rather than five. Again the assumptions made by Hotelling apply with emphasis on the fact that the distribution of grade in the orebody is homogenous and that extraction costs are constant.

The method suggested by Conrad (1999) for the optimal depletion through numerical allocation problem has been used. In this particular case the Solver function in Excel was applied with the target cell being the total of the net profit over the life of the mine (Table II, Column 8) and the constraints being that the discounted value of the rent must be the same in every period, that the sum of the ore produced may not exceed 2500 t, and that the remaining reserve in the last period is zero. Excel Solver only allows for 100 iterations and is not able to resolve the optimization so there is a residual amount of 105.26 t in the last period (Table II, Column 3, Row 8). This last aspect of the constraint is not as rigid as might be hoped, but it nevertheless allows the point to be illustrated.

The following aspects of Hotelling's principle are evident in the depletion of individual mine reserves, as shown in the depletion schedule of Table II and shown graphically in Figure 5.

- The most important aspect is that the rent derived from the depletion of the orebody increases exponentially with time and the value of the discounted rent is the same in every period
- The rate of production from the mine declines over the life of the mine
- The total ore reserve is exhausted in the last period.

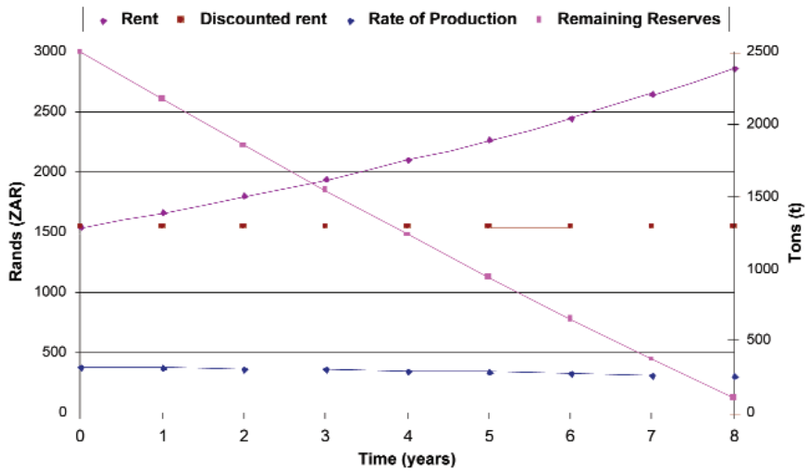

Figure 5—Depletion of the individual mine over eight periods showing rent increasing at the rate of interest and rate of production declining over time

In this example the price of the commodity remains constant throughout the life of the mine, but the average costs decline during the same period, and hence the annual revenue and net profit increase in each period (Table II).

Conclusions

The work of Harold Hotelling in 1931 is considered by many to be the single most important contribution to the understanding of the economics of exhaustible resources ever made. There can be no doubt that at a theoretical level that this is so and it provided the seed for huge numbers of research efforts made during the period 1972 to the early 1990s, nearly two decades when the growing scarcity of natural resources was of concern to many. In terms of a contribution to mining operations the theory describes exactly how a mineral resource should be depleted, provided we can make some strong assumptions.

The experience of extracting mineral resources in the mining industry is, however, significantly different to the proposed theory. Perhaps the single most important contribution made by Hotelling in guiding the schedule of extraction is that the discounted rent derived from the sale of the minerals, should be the same in every period. In order to achieve this end, the single greatest hurdle is that the rate of depletion should decline over the life of the mine, beginning at a maximum and ending at a minimum just before the mine closes.

At the industry level this means that the last ton of ore is being hoisted just as the price of the commodity reaches either a point where it is replaced by the next most inferior orebody, or the choke price, or the commodity is replaced by a backstop technology. At the individual mine level it means that the rate of depletion has moved steadily down the marginal cost curve, from a point where marginal costs equal marginal revenue, to a point where marginal costs equal average costs. At this point the mine should be totally depleted. If not it may continue to operate but it will generate no rent, only normal profit, and can just survive.

These conclusions do not advance our thinking concerning the economics of mining, but it is essential that we explore the economics of exhaustible resources, particularly because it promises so much, but does not deliver much in the way of valuable application in mining. In addition, the simple volume of research demands that we examine the theory and try to understand what it is about the work that has made it so appealing. The continuing search for Hotelling's scarcity rents has a quality about it something akin to the fox-hunt: one is not sure which way it could go next. By and large the average mine operator is not even aware that Hotelling-type scarcity rents exist, far less spend time identifying and scheduling his output according to them (Tilton, 2003).

References

ARROW, K. and CHANG, S. Optimal pricing, use and exploration of uncertain natural resource stocks, Technical Report No. 31, Department of Economics, Harvard University, 1978.

BRADLEY, P.G. Has the 'Economics of Exhaustible Resources' advanced the economics of mining?, Anthony Scott, (ed.): *Progress in Natural Resource Economics Essays in Resource Analysis* by members of the Programme in natural Resource Economics (PNRE) at the University of British Columbia. Clarendon Press, 1985. pp. 317–333.

CARLISLE, D. The Economics of a Fund Resource with Particular Reference to Mine, *American Economic Review*, September 1954, pp. 595–616.

CONRAD, J.M. *Resource Economics, Cornell University*. Cambridge University Press. 1999. 213 pp.

CROWSON, P. The International Tine Agreements: Price stabilisation in action. P. Crowson, *ASTRIDE MINING, Issues and Policies for the Minerals Industry*. Mining Journal Books Ltd, London. 2003. 287 pp.

DASGUPTA, P.S. and HEAL, G.M. The optimal depletion of exhaustible resources, *Review of Economic Studies*, Symposium, 1974. pp. 3–28.

DEVARAJAN, S. and FISHER, A.C. Exploration and scarcity, CMR Working Paper No. IP-290, University of California, Berkley, 1980.

DEVARAJAN, S. and FISHER, A.C. Hotelling's 'Economics of Exhaustible Resources': Fifty Years Later. *Journal of Economic Literature*, vol. XIX, March 1981. pp. 65–73.

EGGERT, R. Personnel communication. 2007.

GRAY, L.C. Rent under the assumption of exhaustibility. *Quarterly Journal of Economics*, vol. 28, May, 1914. pp. 466–489.

HARTWICK, J.M. and OLEWILER, N.D. *The Economics of Natural Resource Use*. Harper and Row Publishers, New York: 1986. 530 pp.

HARTWICK, J.M. and OLEWILER, N.D. *The Economics of Natural Resource Use*, 2nd edition, Addison Wesley Publishers, 1998.

HEAL, G. M. The relationship between price and extraction cost for a resource with a backstop technology. *Bell Journal of Economics*, vol. 7, no. 2. Autumn 1976. pp. 371–378.

HERFINDAHL, O.C. Depletion and the Economic Theory. Mason Gaffney (ed.), *Extractive resources and taxation: Proceedings*. Committee on Taxation, Resources and Economic Development. Madison, University of Wisconsin Press, 1967.

HOEL, M. Resource extraction, uncertainty and learning, *Bell Journal of Economics*, vol. 9, no. 2, Autumn, 1978: pp. 642–645.

HOTELLING, H. The Economics of Exhaustible Resources. *Journal of Political Economy*, vol. 39, April 1931: pp. 137–175.

LEVHARI, D. and PINDYCK, R.C. The pricing of durable exhaustible resources. M.I.T. Energy Laboratory Working Paper EL79-053WP, 1979.

LEWIS, T.R. Attitudes towards risk and the optimal exploitation of an exhaustible resource, *Journal of Environmental and Economic Management*, vol. 4, no. 2, June 1977. pp. 111–119.

LIU, H.C.K. 2002. BANKING BUNKUM, Part 3c: Still more on the US experience. In: Global Economy, *Online Asia Times*, December 10th, 2002. Found at website http://www.atimes.com/atimes/Global_Economy/DL10Dj01.html, 17th July 2007.

LONG, N.V. Resource extraction under the uncertainty about possible nationalisation, *Journal of Economic Theory*, vol. 10, no. 1, February, 1975. pp. 42–53.

MEADOWS AND MEADOWS. The LIMITS to growth: A report for the Club of Rome's project on the predicament of mankind, London, Earth Island, 1972.

NORDHAUS, W.D. The Allocation of Energy Resources, Brookings Papers on Economic Activity, 3, 1973.

PAPPAS, J.L. and HIRSCHEY, M. *Fundamentals of Managerial Economics*. Holt-Saunders International Editors. The Dryden Press, New York: 1985. 586 pp.

PINDYCK, R.S. Optimal exploration and production of non-renewable resources. *Journal of Political Economy* vol. 86, no. 5, October 1978, pp. 841–861.

PINDYCK, R.S. Uncertainty and the pricing of exhaustible resources, M.I.T. Energy Laboratory Working Paper EL79-021WP, 1979.

RICARDO, D. *On the principles of Political Economy and Taxation*. Third Edition, Published by John Murray, London, 1821.

SESINK-CLEE, R. Pay Limits and Cut-off Grades. *Mine Valuation and Grade Control School*, 6 and 7 June 1991, Mintek, Randburg. South African Institute of Mining and Metallurgy, 1991. 24 pp.

SOLOW, R.M. The economics of resources or the resources of economics. *American Economic Review*, vol. 64, no. 2, May 1974, pp. 1–14.

SOLOW, R. M. and WAN, F. Y. Extraction costs in the theory of exhaustible resources. *Bell Journal of Economics, Management. Science*. vol. 7, issue 2. 1976. pp. 359–370.

TIETENBERG, T. H. *Environmental and natural resource economics*. Third Edition. Harper Collins. 1992. 678 pp.

TILTON J.E. On Borrowed Time? Assessing the Threat of Mineral Depletion. Resources for the Future, Washington, DC. 2003. 158 pp. www.rff.org.

WEINSTEIN, M. C. and ZECKHAUSER, R.J. The optimal consumption of depletable natural resources, *Quarterly Journal of Economics*, vol. 89, no. 3, August 1975, pp. 371–392.

WEITZMAN, M. L. The optimal development of resource pools. *Journal of Economic Theory*, vol. 12, no. 3. June 1976. pp. 351–364. ◆

Appendix 1

An analysis of the price paths extraction costs and mineral rents at the industry level

Theoretical background

In a competitive industry where each firm is a price taker and with the market price and rate of extraction connected through the demand curve, each firm sees the industry price rising along a predicted future industry price schedule as it extracts less and less from its stock period by period. The central prediction in exhaustible resource theory is that due to market forces, price will rise along a stable demand curve because of increasing scarcity. In effect the predicted price path is the observed future price path as it could be with a hedging programme.

The industry objective to maximize the present value of rents (NPV) from the depletion of a mineral resource, depends on the choice of a mining rate $(R)t$, over the life of the mine ($t = 0$ to $t = T$), subject to the constraint that the sum of extraction in all periods equals the stock of ore in the ground. There is a so-called opportunity cost—rents (plus their interest) that could have been earned if extraction had occurred in the first period—is associated with extraction in each subsequent period, so rents must be discounted at the rate of interest.

Maximization requires that the cash flows $C(q)$ in each period of production (q tons) be determined. Net Present value in period 1 will be different from that in period 2 because production has changed the size of the remaining ore reserve and hence the value of the constraint. The change in the ore reserve with time, $dS/dt = -q(t)$ and the stock constraint is S, i.e. consumption cannot exceed the original stock S. A simple example of a constrained optimization problem that allows the rent to be maximised over two periods, subject to the constraint that total extraction equals the stock in the ground, is shown by Hartwick and Olewiler (1998) using the method of Lagrangian multipliers.

The objective function is given by

$$\text{Maximise } \pi = \left[B(q_0) - cq_0\right] + \frac{\left[B(q_1) - cq_1\right]}{(1+r)} + \ldots + \frac{\left[B(q_n) - cq_n\right]}{(1+r)^n}$$

Subject to the constraint that the sum of the resources extracted in each period q_0, q_1, etc., is equal to the total stock in the ground S_0 i.e. $q_0 + q_1 = S_0$, where $\pi(q_t)$ is the net benefit from sale of the mineral in period t, q_t is the quantity extracted in time t, c is the constant unit extraction cost and r is the rate of interest.

The constraint is rewritten as $S_0 - q_0 - q_1 = 0$ and the Lagrangian multiplier is introduced so that $\lambda(S_0 - q_0 - q_1) = 0$. This is added to the objective function and set equal to zero, so that the value of the objective function remains unchanged, so we can write a current valued Hamiltonian equation for the problem.

$$L = \left[B(q_0) - cq_0\right] + \frac{\left[B(q_1) - cq_1\right]}{(1+r)} + \lambda(S_0 - q_0 - q_1) = 0 \qquad [1]$$

where λ is the Lagrange multiplier associated with *S*. λ is referred to as the scarcity rent or Hotelling rent. It measures the opportunity cost of foregoing future consumption by extracting one marginal unit of the ore reserves at the current date. Because you mine the marginal unit (one ton or the smallest mining unit) today, you cannot mine it tomorrow, i.e. you have foregone the possibility of mining that unit in future, because you mine it now. Because you have given up the choice of mining it at a later date by mining it now, you incur the so-called opportunity cost or scarcity rent.

The first order conditions for the solution to the maximization of the above equation are derived by differentiating the above equation wrt q_0 q_1 and λ, as follows:

$$\frac{\partial L}{\partial q_0} = p_0 - c - \lambda = 0 \qquad [2]$$

$$\frac{\partial L}{\partial q_1} = \frac{(p_1 - c)}{(1+r)} - \lambda = 0 \qquad [3]$$

$$\frac{\partial L}{\partial \lambda} = S_0 - q_0 - q_1 = 0 \qquad [4]$$

The question is what is the optimal rate of production, i.e. what are the optimal values for q_0 and q_1 over these two periods? By eliminating λ from equations [2] and [3], we get:

$$p_0 - c = \lambda \qquad [5]$$

$$\frac{(p_1 - c)}{(1+r)} = \lambda \qquad [6]$$

so that

$$p_0 - c = \frac{(p_1 - c)}{(1+r)} \qquad [7]$$

This says that the royalty or rent of an exhaustible resource must rise at a rate equal to the social utility rate or the rate of interest *r*, if the social value of the resource is to be maximized. Integrating this equation, we get the famous Hotelling rule that simply says that the price of an exhaustible resource must grow at a rate equal to the rate of interest

$$p_t = p_o e^{rt}$$

where p_0 = price in the initial period and p_t = price in time *t*. The interpretation of the rule is simple. Ore reserves represent a capital asset that is costlessly stored beneath the earth's surface. If they are held without being exploited, their return must be equivalent to the return of other assets and the reserves must produce a capital gain at least at the rate of interest.

The interpretation of λ in Equation [5] is that it is the difference between the present value of price and the marginal cost of extraction, which is the marginal rent or the value of the rent on a single unit of extracted ore. In addition, Equation [7], which is the numerical expression of the flow condition, tells us that the present value of the rent must be the same in each period the mine operates in order to maximize the profits. The identical demand curve example used by Hartwick and Olewiler (1998), namely $p = 700 - 0.25\,q$ is used here, but the example they use is extended to demonstrate the use of the Lagrangian multiplier approach to optimization. This means the choke price is R700. In this example the total stock in the ground to be extracted is also 2500 t, the interest rate is 5 per cent, and R200 is the cost of extraction so that in time period = 0, the net price $p_0 - c$ = (700–0.25 q_0) –200 = 500–0.25 q_0.

By making appropriate substitutions in the objective function, we get:

$$L = [500 - 0.25q_0] + \frac{[500 - 0.25_{q1}]}{(1+0.05)} + \lambda(2500 - q_0 - q_1) = 0 \qquad [8]$$

Differentiating this equation yields three equations as follows:

$$\frac{\partial L}{\partial q_0} = 500 - 0.25q_0 - 0q_1 - \lambda = 0 \qquad [9]$$

$$\frac{\partial L}{\partial q_1} = 476.2 - 0q_0 - 0.238q_1 - \lambda = 0 \qquad [10]$$

$$\frac{\partial L}{\partial \lambda} = 2500 - q_0 - q_1 - 0 = 0 \qquad [11]$$

Rearranging these equations gives

$$0.25q_0 + 0q_1 + \lambda = 500 \qquad [12]$$

$$0q_0 - 0.238q_1 + \lambda = 476.2 \qquad [13]$$

$$q_0 + q_1 + 0 = 0 \qquad [14]$$

And solving them gives: q_0 = 1268.03t, q_1 = 1231.97*t* and λ = 182.99.

Example of a five period optimization for an exhaustible resource

The following example provides an optimization profile for the depletion of 2500 t of ore over five consecutive periods. The marginal cost of extraction is R200/t, the discount rate is 5 per cent and the linear demand curve for the commodity is p = 700 – 0.25 *q*. Differentiation of the objective function after making appropriate substitutions the sets of equations shown in Tables A1 to A4, with corresponding unknowns, is created. At the end of each optimization exercise a value of λ is obtained, but in order to obtain λ for the next period the optimization has to be repeated using the remaining stock in the ground. Thus the value for λ, the undiscounted mineral rent, changes in each period of time.

Table A1

Optimization for the depletion of 2500 t over five consecutive periods

Five periods	Simultaneous equations	Stock extracted
Period 1	$0{,}25q_0 + 0q_1 + 0q_2 + 0q_3 + 0q_4 + 1\lambda$	= 500
Period 2	$0q_0 + 0{,}2381q_1 + 0q_2 + 0q_3 + 0q_4 + 1\lambda$	= 476,2
Period 3	$0q_0 + 0q_1 + 0{,}2268q_2 + 0q_3 + 0q_4 + 1\lambda$	= 453,5
Period 4	$0q_0 + 0q_1 + 0q_2 + 0{,}2160q_3 + 0q_4 + 1\lambda$	= 431,92
Period 5	$0q_0 + 0q_1 + 0q_2 + 0q_3 + 0{,}2057q_4 + 1\lambda$	= 411,35
Stock C	$1q_0 + 1q_1 + 1q_2 + 1q_3 + 1q_4 + 0\lambda$	= 2500

Table A2

Optimization for the depletion of 2500 t over four consecutive periods

Four periods	Simultaneous equations	Stock extracted
Period 1	$0{,}2500q_0 + 0q_1 + 0q_2 + 0q_3 + 1\lambda$	= 500
Period 2	$0q_0 + 0{,}2381q_1 + 0q_2 + 0q_3 + 1\lambda$	= 476.2
Period 3	$0q_0 + 0q_1 + 0{,}2268q_2 + 0q_3 + 1\lambda$	= 453.5
Period 4	$0q_0 + 0q_1 + 0q_2 + 0{,}2160q_3 + 1\lambda$	= 431.92
Stock C	$1q_0 + 1q_1 + 1q_2 + 1q_3 + 0\lambda$	= 1857.26

Value for λ_3 = 356.29 and q_0 = 574.85 q_1 = 503.63 q_2 = 428.63 q_3 = 350.15

Table A3

Optimization for the depletion of 2500 t over three consecutive periods

Three periods	Simultaneous equations	Stock extracted
Period 1	$0{,}2500q_0 + 0q_1 + 0q_2 + 1\lambda$	= 500
Period 2	$0q_0 + 0{,}2381q_1 + 0q_2 + 1\lambda$	= 476.2
Period 3	$0q_0 + 0q_1 + 0{,}2668q_2 + 1\lambda$	= 453.5
Stock C	$1q_0 + 1q_1 + 1q_2 + 0\lambda$	= 1282.41

Value for λ_2 = 374.11 and q_0 = 503.57 q_1 = 428.78 q_2 = 350.06

Table A4

Optimization for the depletion of 2500 t over two consecutive periods

Two periods	Simultaneous equations	Stock extracted
Period 1	$0{,}2500q_0 + 0q_1 + 1\lambda$	= 500
Period 2	$0q_0 + 0{,}2381q_1 + 1\lambda$	= 476.2
Stock C	$1q_0 + 1q_1 + 0\lambda$	= 778.84

Value for λ_2 = 392.83 and q_0 = 428.69 q_1 = 350.15

The results of the analyses shown in Table A1 to A4 are summarized in Table I and are graphically shown in Figure 4 of the main text

The solutions to the equations in Table A1 provide the efficient extraction profile for a non-renewable resource over 5 periods and the value for λ_4 = 339.32:

q_0 = 642.74 q_1 = 574.9 q_2 = 503.46 q_3 = 428.72 q_4 = 350.19 and the sum of the stock extracted is 2500 t.

Appendix 2

An analysis of the rates of production and optimal net benefits (mineral rents) at the individual mine level

The point that the rate of production that maximizes net value per year is different from the rate of production that maximizes the present value over the life of the mine, because of the exhaustible nature of the resource and scarcity, is illustrated using a cost curve from Pappas and Hirschey (1985) and was also investigated by Sesink-Clee (1991).

Total, average, and marginal costs

Short-run profit maximization (optimization) for a mine whose operating environment is bounded by Hotelling assumptions is shown below. We assume fixed costs are fixed and are represented by the point of intersection on the y axis, and the total cost function is a z-shaped curve like that shown in Figure A1. Assume a commodity price of R15 per ton: so total revenue (T*R*), T*R* = 15**q*.

Also, assume a total cost function for the mine T*C* = 15 + $8q - 2q^2 + 1/3q^3$ where R15 is the fixed costs (sunk costs), and is the component of total costs associated with the use of fixed inputs in the short run. A curve showing the behaviour of total costs, fixed costs and variable cost and total revenue is shown in Figure A1. Total fixed costs are constant, even if the firm decides to shut down in the short run it will still have incurred the fixed cost.

Total variable cost is associated with the use of variable inputs in the short run. If total variable costs equal zero it means there are no variable inputs are employed, and TVC will increase as inputs increase. Wages are typically part of the mines total variable costs. If the mine wants to cut back

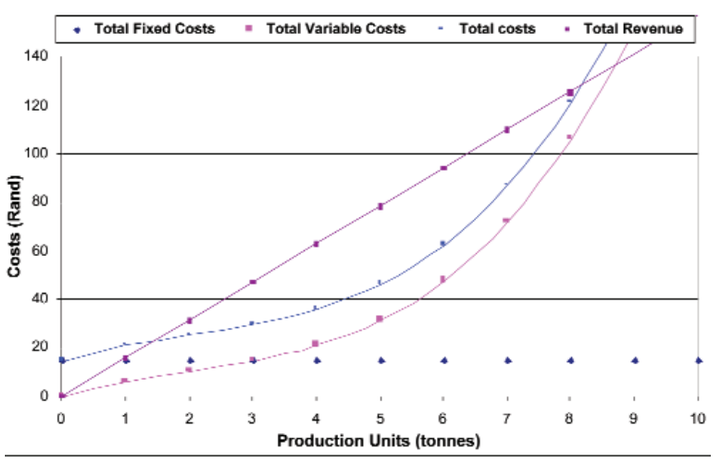

Figure A1—Total fixed costs, total variable costs, total costs and total revenue for a total cost equation TC = 15 + $8q - 2q^2 + 1/3q^3$

production, it can lay off workers, with an associated reduction in wages and hence total variable costs. Total variable costs are given by $TVC = 8q - 2q^2 + 1/3q^3$ (TVC = TC – Fixed costs). Total variable costs are also zero if q equals zero. Average variable cost (AVC) is the variable cost per unit of mine production; so

$$AVC = TVC/q = \left(8q - 2q^2 + 1/3q^3\right)/q$$

$$AVC = 8 - 2q + 1/3q^2$$

Typically average variable costs will initially decrease and then increase with rising production. The minimum point on the average variable cost curve is equal to the derivative of the average variable cost function with respect to q, setting it equal to zero and solving for q.

$$\frac{dAVC}{dq} = -2 + \frac{2}{3}q = 0 \qquad [15]$$

or $q = 3$

A check of the second derivative $d^2AVC/dq^2 = 0{,}666$, confirms that this is a minimum. Average fixed costs (AFC), is fixed cost per unit of output. In this example AFC = $15/q$ so average fixed costs will decrease as production increases. Average total cost (ATC), the total cost per unit of output is the total cost divided by the quantity of production. Average total cost is average cost plus average fixed cost; so ATC = AFC + AVC = $15/q + 8 - 2q + 1/3q^2$

As with the average variable cost, average total cost will decrease and then begin to increase. The minimum point on the ATC curve is also found using the derivative of ATC with respect to q, setting it equal to zero and solving. So

$$\frac{dATC}{dq} = -\frac{15}{q^2} - 2 + \frac{2}{3}q \qquad [16]$$

Setting Equation [16] equal to zero and solving for q gives $2q^3 - 6q^2 - 45 = 45$ and as a solution one of the roots is $q = 4.247$. A check of the second derivative indicates whether $q = 4.247$ is a maximum or a minimum point. The second derivative is

$$d^2ATC/dq^2 = 6q^2 - 12q \qquad [17]$$

Substituting $q = 4.247$ into the second derivative yields

$$d^2ATC/dq^2 = 57.38 \qquad [18]$$

Second derivative greater than zero, confirms that $q = 4{,}247$ is a minimum point. Marginal cost (MC) is the change in total cost that occurs when one additional unit of output is produced (the cost of an additional unit of output). It is the slope of the total cost function, or the first derivative of the total cost function with respect to q.

$$MC = \frac{dTC}{dq} = 8 - 4q + q^2 \qquad [19]$$

Table A5 summarizes the cost information namely TFC, TVC, TC, AFC, AVC, ATC, MC for the total cost function TC = $15 + 8q - 2q^2 + 1/3q^3$ for various levels of output (q). These data are shown graphically in Figure A2.

Lines representing total revenue (TR) = $15q$ and marginal revenue dTR/dQ = MR = 15 are shown in Figure A2. Important relationships evident in Figure A2 include:

- Points 1 and 2 are the break-even levels of production. They are the points where the total revenue curve (TR) intersects the total cost curve (TC) and where the marginal revenue curve (MR) intersects the average total cost curve (ATC).
- Point 3 is the minimum point on the marginal cost curve, MC. At this point, $q = 2$ the first derivative of the marginal cost curve is a minimum point. In this case, because $d = b^2 - 4ac<0$, we have two imaginary roots μ, β as a solution to the marginal cost equation MC = $8 - 4q + q^2$, ($\mu = 2 + 2I$; $\beta = 2 - 2i$). This is also where the second derivative of the total cost curve is zero, $d^2TC/dq^2 = -4 + 2q = 2q = 0$. The inflection point on the total cost curve occurs where $q = 2$, breaking the curve into zones of increasing and diminishing returns. Marginal costs are a minimum at this point.
- Point 4 is the intersection between the marginal cost (MC) and marginal revenue (MR) curves at which MC = MR. At this level of production the slope of the total

Table A5

Various costs for different level of production for the total cost equation TC = 15 + 8q – 2q² + 1/3q³

q	TFC	TVC	TC	AFC	AVC	ATC	MC	Profit
Tons	R	R	R	R/t	R/t	R/t	R/t	
0	15	0	15.0	-	-	-	-	-15
1	15	6.33	21.33	15.0	6.33	21.33	5	-6.33
2	15	10.67	25.67	7.5	5.33	12.83	4	4.33
3	15	15.0	30.00	5.0	5.00	10.00	5	15.01
4	15	21.33	36.33	3.75	5.33	9.08	8	23.67
5	15	31.67	46.67	3.0	6.33	9.33	13	28.34
6	15	48.0	63.00	2.5	8.00	10.50	20	27.01
7	15	72.33	87.33	2.14	10.33	12.48	29	17.68
8	15	106.67	121.66	1.88	13.33	15.21	40	-1.65
9	15	153.0	168.00	1.67	17.00	18.67	53	-32.98
10	15	212.33	228.33	1.50	21.33	22.83	68	-78.3

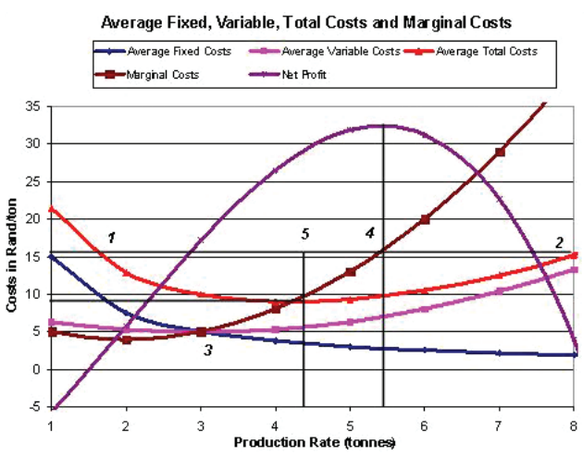

Figure A2—A plot of average fixed, variable, total costs and marginal costs cost functions versus levels of production (q) for the equation TC = 15 + 8q - 2q² + 1/3q³

revenue curve (TR) is parallel to the total cost (TC) curve. This rate of production maximizes net revenue (TR–TC) per year. At this level the owner of a firm in a perfectly competitive market maximizes his profit and the aim of this exercise was to maximize PV.

- Point 5 is the point at which marginal costs (MC) are equal to the average total costs (ATC). It is also the minimum point on the average total cost curve. To the left of point 5, MC is less than AC, and the marginal unit will pull the average down; to the right of point 5, MC is greater than AC and the marginal unit will pull the AC curve up.

Average costs per ton will increase for a mine with costs of R100/ton, if output is increased by one ton and if costs associated with the additional ton are R110/ton. Likewise, if the cost of the additional ton is only R90/ton, average costs per ton will decrease. Minimum AC is the tangent point on the total cost curve, which is intersected by a ray from the origin.

Knowing that total profit (π) equals total revenue (TR) minus total costs (TC) means that for our equation: $\pi = 15q - (15 + 8q - 2q^2 + 1/3q^3)$, the derivative of total profit with respect to q yields:

$$\frac{d\pi}{dq} = 7 + 4q - q^2 \qquad [20]$$

Setting this equation equal to zero and solving for q gives values of q equal to 5.317 and -1.317, and since it is not possible to have negative production, use $q = 5.317$. The second derivative test

$$\frac{d^2\pi}{dq^2} = 4 - 2q = 4 - 2(5.317) = -6.634 \qquad [21]$$

confirms that $q = 5.317$ maximizes the mines total profit. Substituting $q = 5.317$ into our profit equation, indicates that the total profit is R28.66. We can obtain this answer using the concepts of marginal revenue and marginal costs. Doing the same thing algebraically gives:

$$MR = dTR/dq = 15$$

$$MC = dTC/dq = 8 - 4q + q^2$$

Setting marginal revenue equal to marginal costs MR = MC gives:

$15 = 8 - 4q + q^2$ and solving this equation we again arrive at $q = 5.317$ and -1.317. The mines profit per unit will equal = p – ATC

Profit per unit = $15 - (15/q + 8 - 2q + 1/3q^2)$ Substituting $q = 5.317$ indicates that profit per unit equals R5.39. Multiplying it by the number of units produced (5.317) again gives R28.66 as the figure for maximum total profit. Table A6 shows how profit varies against production rate and this is shown in Figure A2.

If the commodity price fell to R9 per ton, rather than remaining at R15 per ton, the mine would earn zero profit; this is equivalent to saying the mine is earning a normal rate of return. This mine will continue to operate indefinitely. If the mine were to shut down now it would lose any money invested in fixed capital. If the mine can make enough revenue to cover all the variable costs, the firm is better off continuing production in the short run. If the commodity price fell to R4 per ton the mine would shut down in the short run, because here 'marginal revenue equals marginal cost' at $q = 2$ so MR = MC and $4 = 8 - 4q + q^2$

However, profit per unit is -R8.67, and total profit for producing two units is -R17.34.

So far this theory is no different from that of an industrial firm owner who wants to arrive at the optimum rate of production for a plant. His optimal rate occurs at point 4, where $q = 5.32$ and marginal costs equal marginal revenues and TR – TC is maximum. This point maximizes his profit because his plant has access to unlimited resource inputs. The important point is that the production rate at which maximizes his profit (or net value) per unit time, also maximizes total net value and present net value, because the plant (under assumed input conditions) has an unlimited life. This assumption does not hold for a mining operation based on a fixed, exhaustible mineral deposit.

Optimizing the rate of production

Any mine owner, because of the exhaustibility of the mineral resource, has a choice to make. He can mine a block of reserve now and shorten the life of the mine or mine the same block at the end of the operation and extend the life of the mine. Mine life is ultimately limited, but is also variable, depending on the rate of production. The rate of production in each mining period (e.g. one year) is different because the size of the remaining reserve is also different. In the hypothetical case presented here, in which the total tonnage to be extracted is 50 t, the rate of production, which maximizes net value per year (point 4, Figure A2) will not maximize the net value over the life of the mine of its present value equivalent. The optimum rate will be less than this point because there is benefit in mining fewer units of reserve per year at a lower average cost, thereby lengthening the life of the mine.

Table A6

Average total costs, marginal costs, net annual profit and profit over the life of the mine for different rates of production

Rate of production tons q	ATC R/t	MC R/t	Annual profit R	Life of mine (year)	LoM profit (R)
0	8.0	-15.0			
1	21.3	5.0	-5.6	50	-28.2
2	12.8	4.0	5.7	25	14.3
3	10.0	5.0	17.1	17	28.5
4	9.1	8.0	26.5	13	33.1
5	9.3	13.0	31.8	10	31.8
6	10.5	20.0	31.2	8	26.0
7	12.5	29.0	22.6	7	16.1
8	15.2	40.0	3.9	6	2.5
9	18.7	53.0	-26.7	6	-14.8
10	22.8	68.0	-71.3	5	-35.7

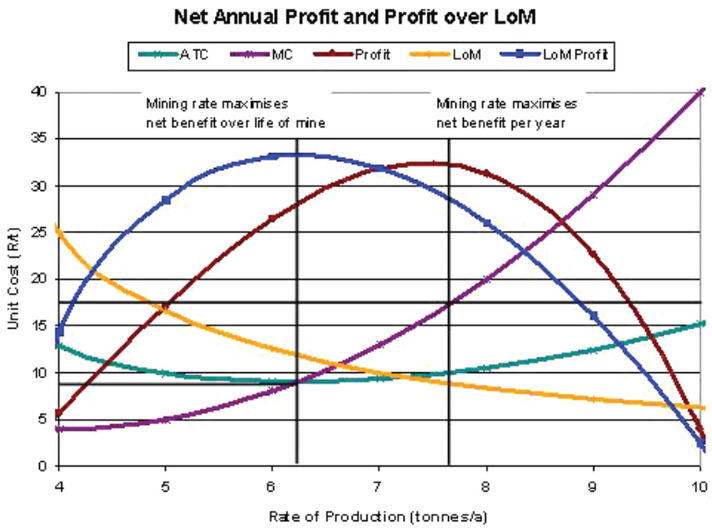

Figure A3—Marginal cost and average cost curves for a mine with a total cost function TC = 15 - 8q - 2q2 + 1/3q3, showing Per Annum Net Profit and Net Profit over the life of the mine

Optimization of the rate of mine production, assuming a total ore reserve of 50 tons, is shown by means of our earlier example, plotted on a slightly larger scale in Figure A3. The two limiting rates of production are point 4, where marginal revenue equals marginal costs and point 5, where average total costs are a minimum.

At point 4 MR = MC the mining rate is 5.35 tons per annum and the life of the mine is 9.35 years.

At point 5 Minimum ATC the mining rate is 4.25 tons per annum and the life of the mine is 11.77 years. The net benefit per year for each of these rates is

- At point 4 5.35 × (15–9.7) = R28.36 per year.
- At point 5 4.25 × (15–9) = R25.50 per year.

This simply confirms what we know already: mining at the rate which marginal cost equals marginal revenue, maximizes net benefits per year. However, if we consider the total net benefits over the life of the mine we have

- At point 4 R28.36 per year for 9.35 years = R265.17
- At point 5 R25.50 per year for 11.77 years = R300.14

Thus mining at the rate were average total costs are a minimum, maximizes total net benefit over the life of the mine.

If time value is considered, the maximum net present value (NPV) occurs somewhere between these two points and it depends critically on the choice of discount rate. As the discount rate *r* increases to infinity, the optimum rate approaches the production rate where MC = MR. As the discount rate falls to zero, the optimum rate of production approaches the minimum AC point.

- As r → ∞, optimum production rate → MC = MR rate.
- As r → 0, optimum production rate → minimum AC rate.

An analysis of the effects of imposing different discount rates of the production rates is presented in Table A7 and shown graphically in Figure A4.

Two clear trends emerge as higher discount rates are applied. The first and obvious trend is that profits either on an annual or life of mine basis are lowered. Secondly the profits calculated on an annual basis fall less rapidly than those determined over the life of the mine. Thirdly the optimal rate of production increased in both cases; and finally the difference in the rate of production for the two cases is reduced.

Furthermore, the optimum rate depends on the size of the ore reserve. When reserves are very high in relation to possible rates of production (e.g. iron ore, coal, and industrial minerals) the optimum rate approaches that for other types of industrial plant.

- As ore reserves → ∞, optimum rate → MC = MR rate.
- As ore reserves → 0, optimum rate → minimum AC rate.

Furthermore, we now show that the optimum rate becomes progressively less as the ore reserves are depleted and the remaining life of the mine shortens.

Table A7

Changes in net profits compared to changes in applied discount rates

Production	0% DCF		5% DCF		12% DCF	
Rate (q)	NAP#	NP LoM*	NAP	NP LoM	NAP	NP LoM
1	-5.6	-28.2	-0.5	-2.5	0.0	0.0
2	5.7	14.3	1.7	4.2	0.1	0.2
3	17.1	28.5	7.6	12.6	0.8	1.4
4	26.5	33.1	14.4	18.0	2.7	3.4
5	31.8	31.8	19.5	19.5	5.1	5.1
6	31.2	26.0	20.8	17.3	6.8	5.7
7	22.6	16.1	15.9	11.4	6.1	4.4
8	3.9	2.5	2.9	1.8	1.3	0.8

#NAP: Net annual profit; ***NP LoM**: Net profit over the life of the mine

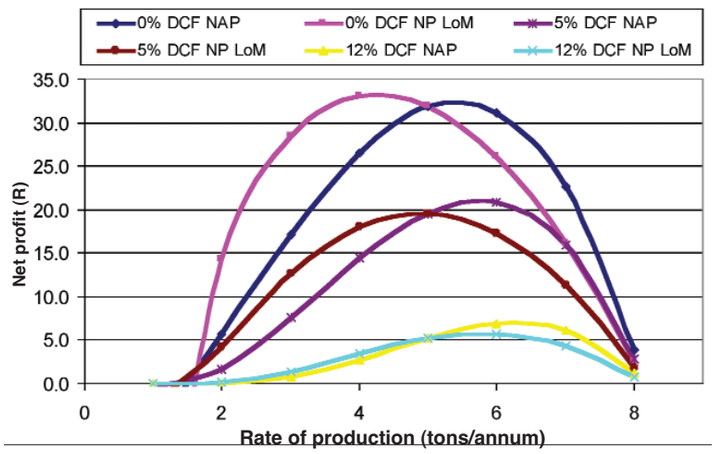

Figure A4—Increasing discount rates leads to reduced profits and increased rates of production

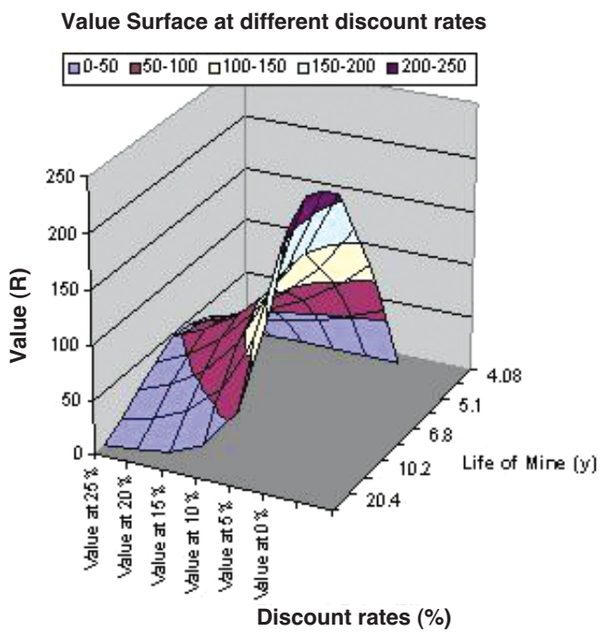

Figure A5—The value surface of a mining operation at different discount rates and life of mine as a proxy for production rates

The value surface

In an extension of this model one might consider the effects of applying different discount rates to the net profit over the life of the mine. This has been done for a range of discount rates from 0% to 25% and the resulting value can now be plotted as a surface that describes the changing value of the deposit with changes in the production rate (*t/a*), the life of mine (y) and the different discount rates as shown in Figure A5.

The value surface describes portion of the topography of a hill that has to be climbed. The surface represents the terrain within which it is possible to move. But rather than relieve oneself of the strain of climbing by making zigzags across the hill, the object of the exercise is to reach the summit of the hill in as short a distance as possible using the steepest portion of the topography as possible. This is in effect what is required if we are to exploit the ore deposit so as to maximize the present value of the operation, but it is contingent on us finding the shortest route to the top of the hill.

Put another way, we can say that the economic contribution of each ton of ore extracted is to be maximized. The three aspects that will affect the topography on the hill are the life of the mine, the discount rates used, and the value of the mining operation, the challenge being to reach the summit with as much value as possible. ◆