
1. Basic economics of nonrenewable resource use

John M. Hartwick

1.1 INTRODUCTION

This is an overview of the economics of extraction of a nonrenewable resource in the tradition of Hotelling (1931). It covers some detail of the structure of the extractive competitive firm and of the extractive industry comprising in one case competitive small firms and in another a monopoly firm. I also present notes on the case of oligopoly in extraction, although this topic is not gone into in detail since it has developed with a variety of approaches and with a large number of interesting contributions. The approach deals with the firm and the industry, and not with economy-wide models. Economy-wide models are well surveyed by van der Ploeg (2011), an article with some 225 references. The overview follows from the first half of a first-year graduate course I have taught over some 20 years. The overview is light on empirical work on the economics of extraction, although there are some brief digressions into this literature. The narrative line is mostly historical.

1.2 L.C. GRAY AND THE EXTRACTING FIRM

Gray (1916) is the first modern economics treatment of an exhaustible-resource problem. He worked with a numerical example in a not transparent fashion and his contribution only became well known in the 1960s (Gaffney 1967). In the meantime the work of H. Hotelling (1931) appeared and received much attention in the 1960s. Gray asks us to consider a competitive firm facing unchanging price p for a unit of extracted material, with homogeneous stock S_0 . There is an extraction cost $C(q_t)$ for $q_t = -\frac{dS_t}{dt}$. His average extraction cost was U-shaped with a marginal extraction cost $C_q(q_t)$ linear in the range he was dealing with. The linear marginal cost specification simplified his calculating. Remember, he was working with a numerical example throughout. The firm is assumed to be a profit maximizer, which for Gray meant selecting an extraction program with the present value of profits from the stream of extractions, q_t , a maximum.

An acceptable program of extractions has $q_0 + q_1 + \dots + q_T = S_0$. For an economist in 1916, this is a very demanding problem to solve since it belongs to the area of dynamic optimization (calculus of variations, dynamic programming, control theory). Gray intuited the correct solution and thus avoided either learning dynamic optimization or expositing his solution with the formalism of dynamic optimization.

Gray worked in discrete time (distinct periods) and intuited that if q_t was the correct extraction for period t , then the maximum present value of profits:

$$\Pi = [pq_0 - C(q_0)] + \frac{1}{1+r} [pq_1 - C(q_1)] + \dots + \left[\frac{1}{1+r} \right]^T [pq_T - C(q_T)]$$

should have a q_{t+1} satisfying:

$$p - C_q(q_t) = \frac{1}{1+r} [p - C_q(q_{t+1})],$$

that is, that marginal profit for date t should be rising at the rate of interest, r for correct extractions, q_t , q_{t+1} , and so on. Alternatively, profit on the marginal unit extracted at each date should be the same, subject to 'normalization' with the interest rate in order to make dollar values in each period the same in present value. Though his costs varied with current quantity extracted, the material being sold at price p was homogeneous and thus should fetch the same value at the margin, over time. In addition Gray intuited that the last extraction in the profit maximizing sequence, namely q_T , should occur where the average cost curve was a minimum. This made marginal profit for his q_T the largest of any in his sequence.¹ Given a correct value for q_T , arriving at the correct sequence of extractions involves working back in time for q_{T-1} , q_{T-2} , to the present or q_0 . Marginal profit shrinks at each date by r percent. His numerical work was fairly imprecise because he was apparently relying on a published table for his present value calculations and the calculations reported in the table were imprecise. He worked with seven periods in his example.

Despite the infelicities, his analysis was quite remarkable – the correct solving of a dynamic optimization problem, well in advance of dynamic optimization becoming a standard tool of analysis. However, he slipped up by invoking an end-point condition that is valid only for solutions of the continuous time version of his problem. In fact the correct discrete-time end-point condition is an approximation to the continuous time version² (Lozada 1993) and so we cannot fault him hugely in this misstep. The principle he arrived at is: the value of the marginal unit (the 'last' unit of the q_t) extracted at each date should be the same, subject to an interest rate adjustment to standardize values for different periods. In addition to

his less than felicitous numerical work in his paper, his verbal rendering of his result was very brief, leaving his complete analysis almost cryptic. When both marginal and average extraction costs are only increasing with q_t , extractive firms end extraction with $q_T = 0$. I will assume such costs are in effect below unless I indicate otherwise.

1.2.1 Value-Change for the Extractive Firm

At date t , the Gray firm will have S_t tons of stock left. This is its asset and the market value of this asset is:

$$V(S_t) = \int_t^{T^*} \pi(q_z^*) e^{-r(z-t)} dz$$

for $\pi(q_t) = pq_t - C(q_t)$ and * indicates a value along the profit maximizing path. What happens to market value as time passes? Since q_t^* is extracted at t , $\pi(q_t^*)$ of value disappears from V_t . In addition the remaining value of stock moves closer to ‘the present’. This shift has value rV_t . Thus:

$$\frac{dV(S_t)}{dt} = -\pi(q_t^*) + rV(S_t).$$

This is not only an accounting result, it is also a calculus result. The value of the firm is shrinking as it depletes its asset, $\frac{dV(S_t)}{dt} = -q_t^* \frac{dV(S_t)}{dS_t}$. This marginal value of the stock at date t is in fact $p - C_q(t)$, current rent on the marginal ton being extracted. This means that current decline in the value of the firm is equal in value to current rent, $q_t^* \frac{dV(S_t)}{dS_t} = q_t^* [p - C_q(t)]$. Samuelson (1964) labels such items as ‘economic depreciation’, that is, the decline in value of an asset when the asset is being used in some sense optimally or efficiently. If the crown or state is deemed to own all subsoil assets, the charge the crown should make to a renter of a mine is the damage the renter does to the crown each period. Since the crown is seeing its asset shrink by $\frac{dV(S_t)}{dt}$ dollars per year, this would seem to be the appropriate ‘royalty’ to ask of the renter, and this value can be expressed as $q_t^* [p - C_q(t)]$, or current ‘total’ rent. Mining economists before Gray reasoned this way about the value of a royalty or user charge. Clearly they were thinking soundly. This value of ‘economic depreciation’ is the one that one inserts into Samuelson’s formula for value invariance of the firm (Samuelson 1964); value invariance of the mining firm under ‘Samuelson taxation’ (see Appendix).

At each instant, then, we have the flow account for the firm:

$$\pi(q_t^*) = rV(S_t) + q_t^* [p - C_q(t)].$$

The decline in the value of the firm along its optimal extraction path is captured by current ‘total’ rent, $q_i^*[p - C_q(t)]$, and current profit from extracting q_i^* divides precisely into current ‘total’ rent plus the time-advancement term, $rV(S_t)$.

1.2.2 The Kuwait Model

A nation living off the sale of a depleting stock such as oil can be thought of as using up the assets that should be gifted ‘in good order’ to the next generation. One can think of the current generation, which is depleting the stock by value $q_i^*[p - C_q(t)]$, deciding to compensate the next generation for the value of depletion. One way to do this is to have the current generation place savings in value $q_i^*[p - C_q(t)]$ into a social responsibility fund, F_t . That is, at each date we have:

$$\frac{dF_t}{dt} = q_i^*[p - C_q(t)]$$

as an investment policy of the current generation. The fund would be yielding the market rate, r . This would leave funds for current consumption, M_t in:

$$\begin{aligned} M_t &= \pi(q_t) - \frac{dF_t}{dt} + rF_t \\ &= rV(S_t) + rF_t. \end{aligned}$$

The time derivative of this is $\frac{dM_t}{dt} = r[\frac{dV(S_t)}{dt} + \frac{dF_t}{dt}]$ which, upon substitution from above, equals zero. Consumption remains constant under the policy of investing $\$X$ in a fund each year, where X equals current ‘economic depreciation’ of the natural capital. This is a variant of the Solow sustainability theorem: investing resource rents in a neoclassical growth model results in a non-varying consumption stream (Solow 1974; Hartwick 1977; Hartwick and Mitra 2013).

1.2.3 Uncertainty Facing the Firm

Suppose our L.C. Gray firm is working with known price p currently and is aware that the price will rise to $p + \alpha$ with probability π at future date T or fall to $p - \beta$ with probability $(1 - \pi)$. It has a known S_0 tons to extract at time $t_0 < T$. To solve this problem we must work back from two alternative end dates. It will extract S^T tons up to date T and $S_0 - S^T$ tons beyond date T , with S^T to be solved for. We choose a value for $S_0 - S^T$ and solve backwards, using the r percent rule, from T^α with the high price

and from T^b with the low price. In solving back with the r percent rule, we obtain two ‘initial’ prices, p^a and p^b in effect just beyond T , indicated by T^+ . Corresponding to these two prices are rents:

$$p^A - C_q(q(T)) = \pi [p^a - C_q(q^a(T^+))] + (1 - \pi) [p^b - C_q(q^b(T^+))].$$

This leaves us to fit in, using the r percent rule, amount S^T between t_0 and T . This exercise gives us a ‘terminal’ rent $p^A - C_q(q(T))$ over the first interval. Our problem in extraction is solved, as S^T is varied and our condition above on rent equality at T is satisfied. Essentially there are then no intertemporal arbitrage profits to be earned as time passes from just before T to just beyond T .

The interesting question is: does uncertainty cause extraction to speed up or slow down? One addresses this question with the Rothschild–Stiglitz (1971) procedure. One perturbs the central first-order condition with respect to the random variable to determine whether the first-order condition is concave or convex in our random price. This is done in Yeung and Hartwick (1988) where it is observed that more uncertainty in price (a mean preserving spread in p) speeds up extraction of the S_0 .

A different problem arises with stock-size uncertainty. Suppose the firm knows that it has 50 tons and may have $50 + \alpha$ tons with probability π or zero more with probability $(1 - \pi)$. To determine the stock size, the firm must first extract the 50 tons. In this case, the date at which uncertainty is resolved depends on how fast the firm extracts its 50 tons. With the r percent rule in operation, this speed, or interval of extraction of the 50 tons, is simply reflected in the firm’s selection of initial quantity, q_0 . Then, given t_0 and q_0 , we get $T(q_0)$ as the date of information resolution and a present value of profit $V(q_0)$. Beyond this date, the firm gets additional discounted profit $V(\alpha)$ with probability π . This $V(\alpha)$ is solved, working back from the end of extraction, with α tons. This backward solving produces an initial quantity q^{T+} . The ‘complementary’ realization yields zero additional profit with probability $(1 - \pi)$. The extraction problem is solved when q_0 maximizes $V(q_0) + e^{-r(T(q_0) - t_0)}\pi V(\alpha)$.

Of interest is the absence of a simple zero profit price arbitrage relationship at $T(q_0)$ in the solution. This is because extraction is yielding both marketable tons and information in the following sense: the date of the arrival of the information depends on the pace of extraction over the first interval. Information has a non-priced value that is ‘distorting’ the ‘traditional’ price arbitrage condition at T , the one which we observed above for the problem of price uncertainty (Hartwick 1983). One can of course apply the Rothschild–Stiglitz procedure to this problem. The result turns out to be that more uncertainty in stock size slows extraction in the first interval.

1.2.4 Exploration and Polluting Extraction

Our extractive firm has been taken as rather pared down, a priori. One extension would have the firm engage in simultaneous extraction and exploration for additional stock. Central here would be the pursuit of current exploration to the point that the value of the marginal ton currently discovered would equal the marginal cost of discovery. This sounds like profit dissipation by excessive exploration but in fact the profit dissipation is only occurring at the margin. There is still ample room for profit intra-marginally from extraction and sale of minerals. There have been observers who have argued that much of the extractive firm's profit will on average be lost to the high cost of exploration activity. This could arise when the exploration is a race by firms over 'open territory'. Claim staking is an institutional mechanism that is designed to mitigate racing or 'gold rush' activity. One can easily imagine that a poorly designed claim-staking procedure could result in excessive or inadequate exploration activity. Pindyk (1978) is an early contribution to the development of a model of the extractive firm, simultaneously engaged in costly exploration activity.

A more classic market failure involving the extractive firm is extraction activity yielding both minerals and pollution simultaneously. Such a view takes us some distance away from classic models of extractive firms. An early investigation is Conrad and Clark (1987, pp. 161–165). See also for example Withagen (1994).

1.3 HOTELLING'S EXTRACTIVE INDUSTRY

Hotelling came to economics with some graduate studies in mathematics and he plunged into extraction theory with a command of the methods of dynamic optimization. He was unaware of Gray's work and turned in extraction economics with a focus on the extractive industry rather than on extractive firms. He worked in continuous rather than discrete time. He left implicit how the industry was divided into extractive firms. He knew that if firms were profit maximizers then one could represent the competitive industry as if it were maximizing the present value of social surplus (current consumer surplus net of current costs of extraction). Hence he sought a path of industry extractions, Q_t , Q_{t+1} , and so on that maximized:

$$B(Q_0) - cQ_0 + \frac{1}{1+r}[B(Q_1) - cQ_1] + \dots + \left[\frac{1}{1+r}\right]^T [B(Q_T) - cQ_T]$$

$$\cong \int_0^T [B(Q_t) - cQ_t] e^{-rt} dt$$

subject to the sum of the Q_t 's remaining less than or equal to the initially endowed stock S_0 . $B(\cdot)$ is the area under a stationary demand schedule and c is the cost of extracting one ton of material (for example, oil).

The solution for a discrete-time version: if Q_t is the correct extraction for period t , then Q_{t+1} should satisfy:

$$\frac{d[B(Q_t) - cQ_t]}{dQ_t} = \left[\frac{1}{1+r} \right] \frac{d[B(Q_{t+1}) - cQ_{t+1}]}{dQ_{t+1}}.$$

This is qualitatively the same condition as we observed for Gray's problem for the extractive firm. Here the value of the marginal ton extracted in periods t and $t + 1$ are the same, subject to an interest rate adjustment that 'normalizes' value across periods. Since marginal consumer surplus, $\frac{dB(Q_t)}{dQ_t}$, is current price p_t , the above condition can be written as:

$$p_t - c = \left[\frac{1}{1+r} \right] [p_{t+1} - c],$$

a condition often referred to as Hotelling's r percent rule for extraction. In continuous time this condition is:

$$\frac{d[p_t - c]}{dt} = [p_t - c]r,$$

a slightly different statement of marginal value increasing at r percent over time. The correct sequence of extractions must satisfy this r percent rule, if the present value of surplus is a maximum.³

In addition, in continuous time, the last quantity extracted in the optimal sequence must satisfy:

$$\frac{d[B(Q_T) - cQ_T]}{dQ_T} Q_T = [B(Q_T) - cQ_T].$$

Given, say, a linear demand schedule, this marginal equals average condition for Q_T implies that $Q_T = 0$ and $\frac{d[B(Q_T) - cQ_T]}{dQ_T}$ is a maximum over all values of the marginal values. For the discrete-time version, the correct end-point condition is an approximation to the continuous time condition.

One way to disaggregate the industry output Q_t to current outputs of firms is to assume that each firm is identical with unit extraction cost c and possessing a very small amount of Q_t a priori. Then as time passes, live firms 'peel off' and 'disappear' as they bring their small holding to market at one date. In this view firm i is indifferent when it extracts because its present value of profit is $p_0 - c$, regardless of when it extracts. A different approach is to view industry output at date t as being produced by a fixed number of identical firms of the L.C. Gray characteristics. Then

each firm's output declines smoothly as time passes. Eswaren et al. (1983) pointed out that in such a specification, notably with each firm with a U-shaped average cost curve for extraction, final industry extraction will occur with a positive Q_T . This implies that the final industry price $p(Q_T)$ will be below the choke price (with say a demand schedule with a finite intercept). With perfect foresight, each firm will have looked forward and will have realized that there is a prospective price jump for output, just beyond date T . Such a prospective price jump is of course associated with a prospective profit jump for a firm with some stock remaining to sell. Hence each firm will be induced to hoard some stock surreptitiously early on. But such behavior means that the so-called dynamic equilibrium solution will not be 'followed' by rational firms. Hence a market failure. Hence extraction costs should not be specified with a declining portion of average extraction costs.

1.4 MONOPOLY EXTRACTION

Hotelling (1931) reported on the case of a single seller doing extraction for the whole industry. The seller can control the price by varying the Q_t brought to market at date t . An intuition here is that in order to stick demanders with relatively high prices, the seller will be obliged to put 'small' Q_t 's on the market at each date. For the same-sized initial stock as the corresponding competitive industry, the average Q_t from the monopolist will be smaller than the corresponding Q_t for the competitive industry. Hence the monopolist will reach exhaustion of their stock over a longer interval or in Hotelling's words: 'the monopolist is the friend of the conservationist'. Hotelling worked with a linear market demand schedule and established his 'friend of conservationist' result. In continuous time the monopolist maximizes:

$$\Pi = \int_0^T Q_t p(Q_t) e^{-rt} dt$$

subject to $\int_0^T Q_t dt = S_0$. The r percent rule governing extraction is:

$$d \left[\frac{d[Q_t p(Q_t)]}{dQ_t} \right] / dt = \frac{d[Q_t p(Q_t)]}{dQ_t} r$$

that is, marginal revenue must be rising at the rate of interest. The end-point condition governing the value of Q_T is:

$$\frac{d[Q_T p(Q_T)]}{dQ_T} = \frac{Q_T p(Q_T)}{Q_T}.$$

Solving involves working back in time, given Q_T , using the r percent rule. Q_0 gets defined by the using up of S_0 over the sequence of Q_t 's.

Sharp-eyed observers (eg. Stigliz 1975) noted that there was a special case for which the monopolist and competitive industry extracted at the same pace: namely when costs of extraction were non-existent and the market demand schedule was of the constant elasticity form. That is, inverse demand is $p_t = A Q_t^{-B}$. In this case marginal revenue is $B p_t$. Hence the percentage change in price is the same as the percentage change in marginal revenue at each instant. A and B are positive constants with $B < 1$. This constant elasticity case figures centrally in the analysis of oligopoly. Eswaran and Lewis (1985) did early and fundamental analysis of oligopoly extraction under the assumption of closed loop competition (Hotelling oligopoly extraction) in part with this demand specification.⁴ See also Benchakroun and Long (2002). Loury (1986) reported on the open loop oligopoly extraction case. Closed loop solutions generally involve solving backward recursions which can become complicated very quickly. See for example Maskin and Newbery (1990). Hartwick and Brolley (2008) investigated closed loop duopoly extraction for the case of a linear market demand schedule and linear extraction costs for each player. They discovered that the closed loop and open loop cases were the same with their specification.⁵ Another form of inter-extractor competition involves a price-taking fringe group competing with a monopoly seller (see the recent contribution of Groot et al. 2003).

Karp and Livernois (1992) investigated the mitigation of monopoly power with a corrective tax. See also Benchakroun and Long (1998).

1.4.1 Technical Progress

The pace of technical progress in extraction of minerals, including oil and natural gas, seems remarkable. Technical progress can work against the current bite of increased scarcity. The slag heaps of years past become places of renewed mining, given higher prices of product and improved recovery technology. The basic Hotelling model predicts more rapid depletion of stock S_0 when unit extraction costs decline (technical progress in extraction). This can be set up as an exercise in comparative statics. One works out the time-path of the extraction of S_0 when extraction costs are c and then re-solves for the extraction path when the extraction costs are $c - \delta$ for $\delta > 0$. One obtains the result: depletion of stock S_0 is sped up when extraction costs decline.

A different comparative statics result also has an important empirical bite. The backstop supply can be thought of as the energy price of the future, the supply activity that will come on stream when our current

low-cost energy source is depleted. If the backstop source gets hit with some technical progress, current suppliers must lower their prices in order to pre-empt an early takeover by the backstop. Hence the scarcity value of the current stock in use reflects more than current demand and stock quantity considerations. Current energy prices for example reflect the expected price at which future technologies can provide energy. When so-called 'cold fusion' was announced as a technical breakthrough in Salt Lake City in 1989, current energy prices 'at the pump' dipped down. They quickly returned to trend when cold fusion was revealed to be incompatible with basic physics.

1.4.2 Empirical Work

Miller and Upton (1985) set up a test of the 'Hotelling valuation principle' (rent on the marginal ton rising at r percent), employing values of extractive firms from the New York stock exchange and other data from the annual reports of their various companies. They reached a qualified positive result; essentially that extractive firms are profit maximizers and the industry is competitive. They located the r percent rule in their data. My personal view is that they did not adequately account for the heterogeneity of the holdings of stock of the various firms (distinct extraction costs for different firms) and I find their results unpersuasive. Cairns and Davis (1998, 2001) worked up their own test of the 'Hotelling valuation principle' with their own dataset and obtained more convincing results, but not strongly supportive of the r percent rule in its simple form. Livernois et al. (2006) searched for Hotelling theory in data on old-growth forests, resources that renew very slowly, with some success. Hartwick et al. (2001) drew on Hotelling theory in an analysis of long-run deforestation, the depletion of a stock.

1.5 STOCKS OF HETEROGENEOUS QUALITIES

Herfindahl (1967) realized that a simple extension of the Hotelling model could be had by starting with distinct deposits with diverse unit extraction costs and diverse sizes. Working back from a terminal Q_T , one would apply the r percent rule to the deposit with the highest unit extraction costs and get an extraction path that ended with the exhaustion of the poorest-quality deposit. Then one would use that 'beginning' price to work back using the r percent rule for the next-highest unit extraction cost deposit, and so on to Q_0 for the lowest-cost deposit, and $p(Q_0)$ at the 'present' date. Such a formulation was pursued in an empirical application by Nordhaus

(1973) in an innovative analysis of the world energy market over the long-run future. Nordhaus implemented the Herfindahl model as a linear program and was able to draw on software to solve easily under differing assumptions about stock availabilities, interest rates, extraction costs, and so on. Nordhaus added geography and transportation costs to his Herfindahl model and ended up with a very complicated 'Hotelling model' as well as some provocative inferences about the way the world energy market was performing. Nordhaus felt that because he did not allow for departures from competitiveness in his formulation, he ended up with a somewhat imperfect simulation of the world energy market.

Nordhaus, in passing, replaced Hotelling's end scenario (price rises to a choke price) with a back stop supply 'mechanism'. In this alternative view, depleting a finite oil stock is not about the world falling off a cliff, rather depleting a finite oil stock is about transitioning to a higher-cost supply, ultimately the 'renewable', high-cost backstop supply. He invoked nuclear fusion power as the 'renewable' backstop supply that world energy markets were headed for. This view leads one to reflect on possible research and development (R&D) scenarios that will see the backstop supply come onstream. Numerous models were subsequently developed that had an early 'Hotelling' phase of energy supply from exhaustible hydrocarbons linked to an end-phase with the backstop supply coming onstream (e.g., Dasgupta et al. 1983; Gallini et al. 1983). Of interest here is how much scope the developers of nuclear fusion power might have to game the 'foreign' suppliers of energy who are relying on depletable stocks.

Kolstad (1994) took up a simpler formulation of Nordhaus's model. He had demanders located in a uniform density on a finite line, with distinct deposits (distinct sizes and distinct unit extraction costs) at each end of the line. Demanders paid a delivery charge per unit, linear in distance. A choke price \bar{p} faced each 'supplier' (competitive firms at each deposit) and the r percent rule governed the time path of rent, λ at each deposit. A complete analysis involves working back for each deposit from a terminal rent and zero quantity, $Q_t = 0$ ($i = L, R$). I illustrate this with an example, with the two deposits of similar sizes and similar unit extraction costs. See Figure 1.1.

Each demander on the line buys one unit per 'period' at the delivered price, provided the delivered price is below the choke price \bar{p} . (Working in continuous time.) We can set the density of sellers at unity. Consider the perfectly symmetric case (unit extraction costs, $c^R = c^L$ and stock sizes (S_0^L and S_0^R) equal and selected so that each market size is initially at $G/2$). I illustrate this in Figure 1.1a. One solves each problem back in time from the top two corners. The r percent rule governs the speed of backward motion. At the end date T , each seller is selling almost zero units at the

amount to R's initial stock and re-solve the supply problem. The extra stock for R (right) means that there will be a new initial interval that has the whole market 'shared' by the two sellers. $p^R(0)$ and $p^L(0)$ are the new initial prices. Consider the solution with time moving forward. As rent, λ rises for each, the mill prices rise at a rate set by r percent. At date T , each delivered price has reached the choke price, \bar{p} . Beyond this date, a hole opens up in the center of the market as each seller's market 'area' shrinks. Over this second phase, each seller's market 'area' declines to zero, not in general reaching zero at the same instant.

One possible scenario in this Kolstad model (similar, not identical, stock sizes for the two deposits and similar, not identical, unit extraction costs) has the low-cost (high-quality) deposit solve with a lower initial rent. This is contrary to possible outcomes for the Herfindahl model (higher-quality deposits extracted from earlier, and higher-quality deposits earn larger initial rents). This seeming anomaly occurs because the high-quality deposit is larger and its relative abundance ends up 'driving down' its initial rent.

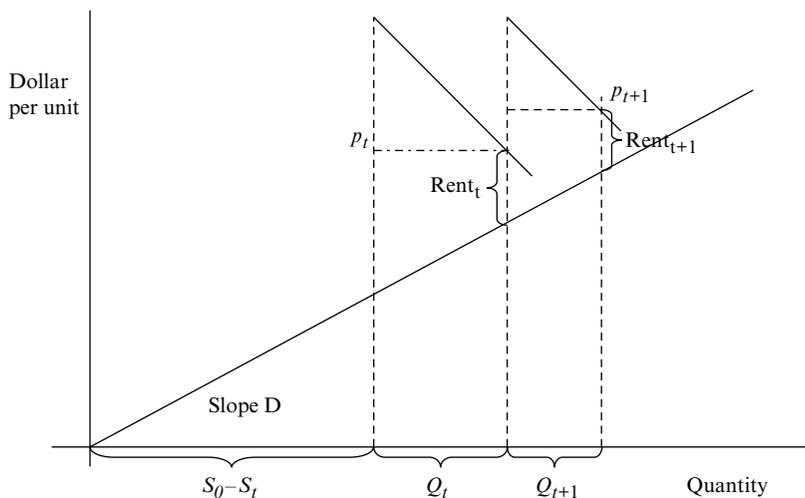
Levhari and Leviatan (1977) presented a smoothly heterogeneous version of the Hotelling model (think of each of Herfindahl's distinct deposits holding a single ton and we now have hundreds of deposits each with a distinct unit extraction cost). The standard formulation assigns an address for where extraction is in the deposit currently. The standard address is 'stock remaining', S_t . Then current extraction cost of Q_t tons is expressed as function $C(Q_t; S_t)$, with $C(\cdot)$ increasing in Q and decreasing in S (larger stock size is associated with lower costs of extracting Q tons). Levhari and Leviatan proceed to work with an industry formulation analogous to that of Hotelling. A question of interest is whether rent on the marginal ton could be declining when mineral quality was posited a priori to be declining over time. See for example Livernois and Martin (2001). Our view is that one should focus closely on the extractive firms, as distinct from industry aggregates, in such an investigation.

In a competitive industry comprising many small firms, each owning one ton, each ton of a different extraction cost, the marginal firm extracting at date t is earning $p_t - c^t$ profit or rent on its ton. To be in a zero profit arbitrage position, it should earn $(1 + r)[p_t - c^t]$ if it delayed extracting by an instant, in say discrete time. Hence an equilibrium path must have p_{t+1} satisfying:

$$p_{t+1} - c^t = (1 + r)[p_t - c^t].$$

This is illustrated in Figure 1.2.

It is also apparent from Figure 1.2 that even though every successive ton extracted is 'higher cost' (poorer quality), rent associated with the marginal



Note: $Q(t)$ in period t and $Q(t+1)$ in period $t+1$, with the marginal firm in period t being the intra marginal firm in period $t+1$. For the marginal firm to be indifferent between extracting between periods, its profit on its ton must differ between periods by $r\%$. Hence $p(t+1) - C(Q(t), S(t)) = (1+r)[p(t) - C(Q(t), S(t))]$.

Figure 1.2 Heterogeneous deposits

ton in each period can be declining as time passes. Detailed examples are easy to construct, using Figure 1.2 as a guide, and continuous-time versions also go through. The inspiration for the model in Figure 1.2 is successive tons being located in a very deep pipe-shaped deposit. Successive tons have the same mineral content, but ‘later’ tons are more expensive to extract than are ‘earlier’ tons.

1.6 DURABLE EXHAUSTIBLE RESOURCES

It is easy to take Hotelling’s 1931 paper as the definitive economics treatment of exhaustible resources. However, implicit in this work was that the resource in question was non-durable in the sense that once a ton was extracted, it was sold and ‘burned up’ completely. There are of course exhaustible resources such as gold, copper, platinum, and so on which are extracted, used in place possibly for many years as roofing material, electrical connectors, and so on, and possibly recycled. Such durable resources may ultimately get ‘burned up’ in a thermodynamic sense but they are also typically long-lived in use before they ‘evaporate’. What is interesting

however is that a Hotelling approach to such resources yields a Hotelling-like rule for extracting them from a deposit. I now turn to this.

Our Hotelling approach involves social welfare maximization (the implicit assumption of competitive suppliers). There is then a stationary demand schedule to quantities currently extracted. However, current value from current extraction must also incorporate past extractions that are still in use. We can simplify and assume once extracted a ton provides undiminishing services forever. Hence if Q_t is currently extracted (with $Q_t = -\frac{dS_t}{dt}$) from S_t , current value (consumer surplus) is $B(Q_t + Q_0 + Q_1 + \dots + Q_{t-1})$. We can take extraction costs as increasing in Q_t in $C(Q_t)$. Hence our 'Hotelling problem' is the choice of Q_0, Q_1, \dots, Q_T to maximize:

$$\begin{aligned}
 W = & [B(Q_0) - C(Q_0)] + \left[\frac{1}{1+r} \right] [B(Q_0 + Q_1) - C(Q_1)] \\
 & + \left[\frac{1}{1+r} \right]^2 [B(Q_0 + Q_1 + Q_2) - C(Q_2)] \\
 & + \left[\frac{1}{1+r} \right]^3 [B(Q_0 + Q_1 + Q_2 + Q_3) - C(Q_3)] \\
 & + \dots + \left[\frac{1}{1+r} \right]^T [B(Q_0 + Q_1 + Q_2 + \dots + Q_T) - C(Q_T)]
 \end{aligned}$$

Subject to $Q_0 + Q_1 + \dots + Q_T = S_0$. The first-order condition (Euler equation) for this problem is:

$$[B_{Z_t}(Z_t) - C_{Q_t}] + \left[\frac{1}{1+r} \right] C_{Q_{t+1}} = 0$$

for $Z_t = Q_0 + Q_1 + \dots + Q_t$. $B_{Z_t}(Z_t)$ is what I referred to as p_t earlier in my analysis of Hotelling (1931). Thus we have $p_t - C_{Q_t}$ is equal to a negative number. The explanation of this anomaly is that p_t is a use-rent per unit for a single period. When one buys a unit of gold, one is acquiring the use of it in perpetuity, that is, over many periods into the future. Let us define:

$$Y_t = p_t + \left[\frac{1}{1+r} \right] p_{t+1} + \left[\frac{1}{1+r} \right]^2 p_{t+2} + \dots - C_{Q_t}$$

and:

$$Y_{t+1} = p_{t+1} + \left[\frac{1}{1+r} \right] p_{t+2} + \left[\frac{1}{1+r} \right]^2 p_{t+3} + \dots - C_{Q_{t+1}}.$$

Then:

$$Y_t - \left[\frac{1}{1+r} \right] Y_{t+1} = p_t - C_{Q_t} + \left[\frac{1}{1+r} \right] C_{Q_{t+1}}.$$

The right-hand side of this is identical to the Euler equation above. Hence we infer that optimal extraction requires that value Y_t rise at the rate of interest. This rule can be written as:

$$V_t - C_{Q_t} = \left[\frac{1}{1+r} \right] [V_{t+1} - C_{Q_{t+1}}]$$

for V_t , the present value of use-rent terms (the p_t 's). Clearly V_t is a capital value or capital price. Hotelling's rule for extraction in his 1931 analysis becomes a special case of this extraction rule, special in the sense that the price in his rule is a one-period price, whereas our V_t is the multi-period analogue of his price. Durability is tricky because in our formulation, p_t , the single-period use-price, will always be declining. We can have $p_t - C_{Q_t} < 0$ for some dates, as we observe is required for all dates above in our Euler equation, and $V_t > 0$ at the same time. But it is plausible that for some demand schedules and cost functions $V_t - C_{Q_t}$ will turn negative or possibly not rise at r percent. We could get around this 'pathology' by introducing explicit decaying or 'rusting' of our extracted material as time passes. With a high enough rate of 'rusting' we can easily get a series with $V_t - C_{Q_t}$ rising at r percent to the end (exhaustion of the stock).

This analysis is indirectly involved with extraction by competitive small firms. Coase (1972) argued that complicated issues of strategy between consumers and the supplier arise when a monopolist controls extraction. See for example Karp (1993) for an analysis.

1.7 END COMMENT

There seems to be a direct link from Hotelling (1931) to Solow (1974) and subsequent work by economists on measuring natural capital and sustainability. Sustainability has become linked to issues in 'the resource curse' and the development of Third World nations. Reflections on the sustainability of a nation's economic well-being has led quite directly to reflections on the issue of sustainability of the Earth's economy. Hotelling's work and derivatives from it provided observers with a way to structure their thinking about measuring natural capital and sustainability (Hartwick 1990; Ruta and Hamilton 2007; Arrow et al. 2012). Precise thinking and good-quality measurement provide a foundation for policy design for our planet. And policy intervention in our use of resources on and in the planet

will not be of less concern to our successors. The analysis of nonrenewable resource use, in the Hotelling tradition, is not a bad way to get better at thinking about the future evolution of our planetary economy.

NOTES

1. The correct end-point condition for the continuous time version is:

$$\frac{d[pq_T - C(q_T)]}{dq_T} = \frac{pq_T - C(q_T)}{q_T}.$$

Marginal profit equals average profit. See Gelfand and Fomin (1963, p. 60).

2. In discrete time the last quantity q_T should be positive, have its marginal profit r percent above that for q_{T-1} , and this marginal profit should be close in value to that for the continuous time q_T .
3. Gaudet (2007) focuses on this rule in an interesting overview of extraction economics. He inquires about how it is altered by certain forms of extraction costs, uncertainty, and market structure. One wants to keep track of quantities in the market at a date in time and quantities being managed by representative firms at the same date in this interesting approach. Gaudet is attempting in part to take 'Hotelling theory to the data' in his analysis.
4. See also Benchakroun and Long (2002). Loury (1986) reported on the open loop oligopoly extraction case.
5. Van der Ploeg (2010) took up a dynamic oligopoly formulation in an effort to focus on excessively rapid extraction. His analysis turned out not to provide strong support for this intuition.

REFERENCES

- Arrow, Kenneth J., Partha Dasgupta, Lawrence H. Goulder, Kevin J. Mumford and Kirsten Oleson (2012) 'Sustainability and the measurement of wealth', *Environment and Development Economics*, 17 (3), 317–353.
- Benchekroun, Hassan and N.V. Long (1998) 'Efficiency inducing taxation for polluting oligopolists', *Journal of Public Economics*, 70, 325–342.
- Benchekroun, Hassan and Ngo Van Long (2002) 'On the multiplicity of efficiency-inducing tax rules', *Economics Letters*, 76 (3), 331–336.
- Cairns, Robert and Graham Davis (1998) 'On using current information to value hard-rock mineral properties', *Review of Economics and Statistics*, 80 (4), 658–663.
- Cairns, Robert and Graham Davis (2001) 'Adelman's Rule and the petroleum firm', *Energy Journal*, 22 (3), 3–54.
- Coase, Ronald, (1972) 'Durability and monopoly', *Journal of Law and Economics*, April (15), 143–149.
- Conrad, Jon and Colin W. Clark (1987) *Natural Resource Economics: Notes and Problems*, New York: Cambridge University Press.
- Dasgupta, Partha, Richard Gilbert and Joseph Stiglitz (1983) 'Strategic considerations in invention and innovation: the case of natural resources', *Econometrica*, 51 (5), 1439–1448.
- Eswaran, Mukesh, T. Heaps and T.R. Lewis (1983) 'Exhaustible resource markets with decreasing costs', *Journal of Political Economy*, 91, 154–167.
- Eswaran, Mukesh and Tracy R. Lewis (1985) 'Exhaustible resources and alternative equilibrium concepts', *Canadian Journal of Economics*, 18 (3), 459–473.

- Gaffney, Mason (ed.) (1967) *Extractive Resources and Taxation*, Madison, WI: University of Wisconsin Press.
- Gallini, Nancy, Tracy Lewis and Roger Ware (1983) 'Strategic timing and pricing of a substitute in a cartelized resource market', *Canadian Journal of Economics*, 16 (3), 429–446.
- Gaudet, Gérard, (2007) 'Natural resource economics under the rule of Hotelling', *Canadian Journal of Economics*, 40, 1033–1059.
- Gelfand, I.M. and S.V. Fomin (1963) *Calculus of Variations*, Englewood Cliffs, NJ: Prentice-Hall.
- Gray, L.C. (1916) 'Rent under the assumption of exhaustibility', *Quarterly Journal of Economics*, 28, 466–487.
- Groot, Fons, Cees Withagen and Aart de Zeeuw (2003) 'Strong time-consistency in the cartel-versus-fringe model', *Journal of Economic Dynamics and Control*, 28 (2), 287–306.
- Hartwick, John M. (1977) 'Intergenerational equity and the investing of rents from exhaustible resources', *American Economic Review*, 67 (5), 972–974.
- Hartwick, John M. (1983) 'Learning about and exploiting exhaustible resource deposits of uncertain size', *Canadian Journal of Economics*, 16 (3), 391–410.
- Hartwick, John M. (1990) 'Natural resources, national accounting and economic depreciation', *Journal of Public Economics*, 43 (3), 291–304.
- Hartwick, John M. and Michael Brolley (2008) 'The quadratic oil extracting oligopoly', *Resource and Energy Economics*, 30 (4), 568–577.
- Hartwick, John M. and Tapan Mitra (2013) 'New cases of Solow sustainability with exhaustible resources', manuscript.
- Hartwick, John M., Ngo Van Long and Huilan Tian (2001) 'Deforestation and development in a small open economy', *Journal of Environmental Economics and Management*, 41 (3), 235–251.
- Herfindahl, Orris C. (1967) 'Depletion and economic theory', in M. Gaffney (ed.), *Extractive Resources and Taxation*, Madison, WI: University of Wisconsin Press, pp. 63–90.
- Hotelling, Harold (1931) 'The economics of exhaustible resources', *Journal of Political Economy*, 39, 137–175.
- Karp, Larry S. (1993) 'Monopoly extraction of a durable nonrenewable resource: failure of the Coase conjecture', *Economica*, 60 (February), 1–11.
- Karp, Larry and John Livernois (1992) 'On efficiency-inducing taxation for a non-renewable resource monopolist', *Journal of Public Economics*, 49 (2), 219–239.
- Kolstad, Charles D. (1994) 'Hotelling rents in Hotelling space: product differentiation in exhaustible resource markets', *Journal of Environmental Economics and Management*, 26 (2), 163–180.
- Levhari, David and Nissan Liviatan (1977) 'Notes on Hotelling's economics of exhaustible resources', *Canadian Journal of Economics*, 10 (2), 177–192.
- Livernois, John and Patrick Martin (2001) 'Price, scarcity rent, and a modified r per cent rule for non-renewable resources', *Canadian Journal of Economics*, 34 (3), 827–845.
- Livernois, John, Henry Thille and Xianqiang Zhang (2006) 'A test of the Hotelling rule using old-growth timber data', *Canadian Journal of Economics*, 39, 163–168.
- Loury, G. (1986) 'A theory of oiligopoly: Cournot equilibrium in exhaustible resource markets with fixed supplies', *International Economic Review*, 27 (2), 285–301.
- Lozada, Gabriel A. (1993) 'Existence and characterization of discrete-time equilibria in extractive industries', *Resource and Energy Economics*, 15 (3), 249–254.
- Maskin, Eric and David Newbery (1990) 'Disadvantageous oil tariffs and dynamic consistency', *American Economic Review*, 80 (1), March, 143–156.
- Miller, Merton H. and Charles A. Upton (1985) 'A test of the Hotelling valuation principle', *Journal of Political Economy*, 93, 1–25.
- Nordhaus, William D. (1973) 'The allocation of energy resources', *Brookings Papers on Economic Activity*, 3, 529–570.
- Pindyck, Robert S. (1978) 'The optimal exploration and production of nonrenewable resources', *Journal of Political Economy*, 86 (5), 841–861.

- Rothschild, Michael and Joseph E. Stiglitz (1971) 'Increasing risk II: economic consequences', *Journal of Economic Theory*, 3 (1), 66–84.
- Ruta, Giovanni and Kirk Hamilton (2007) *The Capital Approach to Sustainability*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
- Samuelson, Paul A. (1964) 'Tax deductibility of economic depreciation to insure invariant valuations', *Journal of Political Economy*, December, 604–606.
- Solow, Robert M. (1974) 'Intergenerational equity and exhaustible resources', *Review of Economic Studies*, Symposium, pp. 29–46.
- Stiglitz, Joseph E. (1975) 'Monopoly and the rate of extraction of exhaustible resources', *American Economic Review*, 66 (4), 655–661.
- van der Ploeg, Frederick (2010) 'Voracious transformation of a common natural resource into productive capital', *International Economic Review*, 51 (2), 365–381.
- van der Ploeg, Frederick (2011) 'Natural resources: curse or blessing', *Journal of Economic Literature*, 49 (2), 366–420.
- Withagen, Cees A.A.M (1994) 'Pollution and exhaustibility of fossil fuels', *Resource and Energy Economics*, 16, 235–242.
- Yeung, David and John M. Hartwick (1988) 'Interest rate and output price uncertainty and industry equilibrium for non-renewable resource extracting firms', *Resources and Energy*, 10 (1), 1–14.

APPENDIX: TAXATION OF THE FIRM FOR VALUE INVARIANCE

In the absence of taxation we have our change in the value of the extractive firm in:

$$\frac{dV(S_t)}{dt} = -\pi(q_t^*) + rV(S_t)$$

where S_t is stock remaining and q_t^* is current extraction along a profit maximizing path. Samuelson taxation has: (1) the current depreciation allowed for tax purposes equals 'true economic depreciation', $\frac{dV(S_t)}{dt}$, and (2) the current interest on the debt of the firm is deductible (this turns out to be equivalent to the firm borrowing at $r(1 - \tau)$ where τ is the current rate of taxation of profit). Hence under Samuelson taxation, the current taxable profit of the firm is $\pi(q_t) - D_t$ and profit net of tax is $\pi(q_t) - \tau[\pi(q_t) - D_t] = (1 - \tau)\pi(q_t) + \tau D_t$, where D_t is allowable depreciation. The new post-tax value of the firm is:

$$V^\tau(S_t) = \int_t^T [(1 - \tau)\pi(q_z) + \tau D_z] e^{-r(1 - \tau)[z - t]} dz.$$

The time derivative of this is:

$$\frac{dV^\tau(S_t)}{dt} = (1 - \tau)rV^\tau(S_t) - [(1 - \tau)\pi(q_t) + \tau D_t].$$

When D_t is set at $\frac{dV^\tau(S_t)}{dt}$, we have:

$$\frac{dV^\tau(S_t)}{dt} = rV^\tau(S_t) - \pi(q_t)$$

which is the same expression as we had for the no-tax case. Hence under Samuelson taxation, the value of the firm under taxation is the same as under no taxation. This is the 'value invariance' of the Samuelson tax scheme. Note that we have not made essential use of true economic depreciation here being represented by current resource rent. The attractiveness of 'applying' Samuelson's taxation theory to our extractive firm is that we have a detailed understanding of the 'true economic depreciation' that is driving the tax-neutrality result.