

Extraction costs in the theory of exhaustible resources

Robert M. Solow

Professor of Economics
Massachusetts Institute of Technology

and

Frederic Y. Wan

Professor of Mathematics and Director of the Institute of Applied Mathematics and Statistics
University of British Columbia

This paper has two objectives. First, it reformulates the theory of optimal use of an exhaustible resource with more attention to the costs of extraction than has been customary in the literature. The output and shadow-price implications of optimal extraction are studied under these broader assumptions. Second, the paper provides some numerical solutions of a simple two-grade case, to give some feeling for the quantitative importance of changes in the supply of exhaustible resources. Our most striking result is, in fact, the suggestion that relatively large changes in resource availability generate very small changes in the sustainable level of final consumption.

The newer aggregative literature on the socially optimal management of a given pool of exhaustible resources has ignored one important characteristic of natural resources: the differential quality of mineral deposits. This note is intended to fill the gap by offering a treatment of differential extraction costs in the context of an aggregative model. We embed the analysis of differential quality in one particular formulation of the macroeconomic problem of the optimal exploitation of exhaustible resources; but we think it would carry over straightforwardly to other ways of setting up the problem.

The basic economic model is that of Solow (1974), so it need only be sketched here. There is a single produced commodity, whose output (Q) can be either consumed directly (C) or accumulated as a stock of reproducible capital (K). Output is produced under constant returns to scale according to a well-behaved production function $Q = F(K, R, L)$, whose inputs are the services of capital and labor and the using-up of resources. Population and employment (L) are constant, so-given constant returns to scale-we might as well normalize to $L = 1$, after which Q , K , C , and R can be thought of as output, capital, consumption, and resource use per head. R , as mentioned, is the flow of the single natural resource into production. For reasons given in Solow (1974), we take $F(K, R, L)$ to be the Cobb-Douglas:

$$Q = K^a R^b L^{1-a-b}. \quad (1)$$

R is therefore indispensable to production, but there is no limit to the

1. Introduction

2. Formulation of the problem

output that can be produced from a given flow of resource input, if only there are enough labor and capital available.¹ There is no technological progress. (Many of these assumptions are made just to pose the problem of resource exhaustion starkly but not hopelessly.)

The natural resource appears in the production function as a homogeneous flow extracted from a given, nonaugmentable, preexisting stock. If all resource deposits were of the same quality, one would constrain the time-path of the economy by $\int_0^{\infty} R(t)dt \leq \bar{R}$. But resource deposits differ in quality. The differences are summed up in θ , defined as the number of units of producible output used up in extracting one unit of the resource for use in production; θ is a natural measure of cost associated with the use of a particular resource deposit. The resource endowment of the economy at the start of the planning period is described by a density function $f(\theta)$: $f(\theta)d\theta$ is the amount of the resource available at a unit cost between θ and $\theta + d\theta$.² The finiteness of resources at time zero is now expressed by

$$\int_{\theta}^{\bar{\theta}} f(\theta)d\theta = \bar{R}, \quad (2)$$

where $\bar{\theta}$ is the leanest, 'highest-cost, ore available. (Actually it might not do violence to the concept of exhaustibility to let both \bar{R} and $\bar{\theta}$ go to infinity. One could perhaps imagine infinite resource availability at infinite cost. But we do not pursue that line here.)

Now let $\theta(t)$ be the quality of the resource actually mined at time t . (For the time being we shall simply assume that the economy exploits its low-cost deposits first, then high-cost, then higher-cost, and so on. Thus $W(t) > 0$. We shall come back to this question explicitly later on.) Then the connection between $R(t)$ and $\theta(t)$ is given by

$$\int_0^t R(u)du = \int_0^{\theta(t)} f(u)du, \quad (3)$$

which says that the total use of resources in production between the initial time and now is simply that part of the initial resource endowment with quality no worse than what is now being mined. Since this holds for every t , we can differentiate with respect to time to get

$$R(t) = f[\theta(t)]\theta'(t). \quad (4)$$

This equation can be substituted into the production function, so that the whole problem can be analyzed in terms of θ rather than R , if that should be convenient.

The depreciation of reproducible capital is of no particular interest here, so we might as well think of Q as already net of depreciation. The alternative uses of output are then consumption, net investment in reproducible capital, and the cost of extracting the natural resource; thus

$$Q = C + K + \theta R. \quad (5)$$

The operation of the economy is now described by (1), (4), and (5).³ At any instant the economy is characterized by its (constant)

¹ We assume $a > b$ for reasons explained in Solow (1974).

² No difficulty of principle would be involved in allowing the distribution of resources by quality to be discrete or mixed. Later on, we do sample computations with a two-point distribution.

³ From the form of (3) and (5), one can see that the model will bear an alternative

population, the stock of capital it has inherited from the past, the cost parameter describing its cheapest remaining resource deposit, and the distribution of its remaining resource deposits by quality. The economy must choose how much of the natural resource to mine and use currently: so $R(t)$ is a decision variable. Once it is decided, current output is determined by (1). According to (5), an amount of output equal to $Q - \theta R$ is available for consumption and investment. The economy must then decide the allocation of output between those two uses. The allocation of net investment determines the stock of capital for the next instant. Equation (4) shows how far this instant's mining activity pushes forward the cost parameter of the best deposit available for the next instant's mining. The data have then been completely regenerated for repeating the process just described. The optimum problem for society is to make its decisions on the rate of resource depletion and on the division of net output between consumption and investment so that the resulting path through time is the best possible. The possible paths are limited by the finiteness of resources (2), the initial stock of capital goods K_0 , and the fact that the stock of capital can not become negative (though it could be run down toward zero).

To complete the statement of the problem we need a criterion for choosing among possible paths. The conventional choice is, of course, the maximization of a discounted time-additive concave utility function depending only on consumption per head, i.e., $\int_0^{\infty} e^{-\rho t} u[C(t)] dt$. Solow (1974) analyzed such a criterion, but concentrated on what amounts to an interesting limiting case of it: maximization of the permanently sustainable constant level of consumption per head. (This is a direct application of the Rawls max-min criterion to the intertemporal context.) That is to say, one puts $C = \text{constant}$ in (5) and seeks the largest such constant associated with a possible path. The discount rate is zero. We continue with that criterion here, not because it is clearly superior, but because it will suffice for showing what we want to show, namely the implications of the differential quality of resource deposits.

One could make a case for including in the choice criterion a "conservative motive," according to which society would draw direct satisfaction from the remaining stock of the resource, as well as from current consumption. But that too would be a diversion from the main theme. The sort of natural resource to which this model applies is thus a productive input, a mineral fuel like coal or oil, or a metallic mineral like bauxite or zinc, to which the conservation motive does not seem to apply. Under uncertainty about future demands and technology, the current reserve might provide a kind of security-satisfaction, but even that is derived from an underlying valuation defined over realizations of an uncertain consumption process.

interpretation, in which the resource is homogeneous, but the unit cost of extraction is an increasing function of the cumulative amount already extracted. To be precise, let F be the cumulative function corresponding to the density f . Let $S(t) = \int_0^t R(u) du$, i.e., $S(t)$ is cumulative extraction up to t . Then (3) reads $S(t) = F(\theta(t))$ and, since F is nondecreasing, $\theta(t) = F^{-1}(S(t))$, with jumps where F is horizontal. Thus, the cost of extracting R at time t is $\theta R = RF^{-1}(S(t))$. So (5) covers a class of cases in which unit cost of extraction is an increasing function of cumulative extraction to date, but independent of the current flow rate of extraction.

3. A digression on efficiency

In arriving at (4) we have assumed that society uses up its endowment of natural resources strictly in order of increasing extraction cost, so that θ is an increasing function of time. This could come about for strictly physical reasons. The richer ores might be nearer the surface and have to be removed before the lower-lying leaner ores can be used at all. Or else nearness to the surface might be precisely what generates the cost differential in the first place. To recognize that, however, is to recognize that the situation might happen to be reversed, with the richer ores lying beneath the leaner ones. In that case the assumption would be wrong and have to be reversed. In this paper, we want to suppose that the whole endowment of resources can be used in any order and then argue that it is rational to use the rich ores first, under very weak assumptions.

We are going to make this argument in a simpler version of the model already proposed. We shall take it in discrete time; indeed, to start with we assume there are only two periods of time involved. Secondly, we shall temporarily interpret Q as gross output, and regard reproducible capital as circulating capital, so the whole of the capital used in production in period 1 is used up and the whole of the capital to be used in production in period 2 has to be provided out of Q_1 , the output of period 1. This temporary change in assumptions is only for simplicity's sake.

So the situation is this. The economy has a constant population L . It starts with an endowment of reproducible capital K_1 . It also has a quantity \bar{R}_1 of the natural resource with cost parameter θ_1 and an amount \bar{R}_2 at cost θ_2 , where $\theta_2 > \theta_1$. Designate by R_{11} and R_{21} the consumption of low-cost and high-cost resource in period 1 and by R_{12} and R_{22} the analogous quantities in period 2. We then have

$$Q_1 = F(K_1, L, R_{11} + R_{21}) = C + K_2 + \theta_1 R_{11} + \theta_2 R_{21}$$

and

$$Q_2 = F(K_2, L, R_{12} + R_{22}) = C + \theta_1 R_{12} + \theta_2 R_{22}.$$

The resource constraints are

$$R_{11} + R_{12} = \bar{R}_1$$

and

$$R_{21} + R_{22} = \bar{R}_2.$$

We now show that if R_{12} and R_{21} are both positive, the situation cannot be optimal. That is to say, it cannot be optimal to use any of the high-cost resource in period 1 unless the low-cost resource has been exhausted in period 1. The only additional condition needed is that the net marginal product of reproducible capital in period 2 be positive.

Suppose that R_{12} and R_{21} are both positive. Then there is a positive number x sufficiently small that the intertemporal allocation of resources can be changed to $R_{11} + x$ and $R_{21} - x$ in period 1 and $R_{12} - x$ and $R_{22} + x$ in period 2. That is to say, the use of low-cost ore is increased by x in period 1 and decreased by x in period 2, while the use of high-cost ore is reduced by x in period 1 and increased by x in period 2. This clearly leaves the resource constraints satisfied and also leaves the gross outputs Q_1 and Q_2 unchanged. There is a *saving* of extraction costs in period 1 equal to $(\theta_2 - \theta_1)x$. Leave C unchanged

and add the output thus saved to K_2 , making it $K_2 + (\theta_2 - \theta_1)x$. Now in period 2 gross output can be increased, to first order, by $(\theta_2 - \theta_1)x F_{K_2}$ where F_{K_2} is the marginal product of capital in period 2. On the other hand, extraction costs in period 2 are *increased* by $(\theta_2 - \theta_1)x$ because of the substitution of lean for rich ore. The change in the output available for consumption in period 2 is $(F_{K_2} - 1)(\theta_2 - \theta_1)x$ and this is strictly positive provided that F_{K_2} exceeds one. Since this is a circulating-capital model, the condition is only that the net marginal product of reproducible capital be positive. But now it is obvious that the initial allocation was not optimal, because some of the excess consumption capacity in period 2 could be transferred backwards to period 1 (by reducing slightly the increased carry-forward of capital, for instance) so that consumption could be increased in both periods.

Thus, if the economy is not saturated with reproducible capital, it is not optimal to use any high-cost resource until the low-cost resource is exhausted, which is what we set out to prove. Now the argument could proceed in exactly the same way if period 2 were required to leave a fixed terminal capital K_3 out of its gross output. But this means that the same argument would work if periods 1 and 2 were any neighboring periods in a much longer process. So the conclusion is not confined to two-period histories. Indeed, one can see how a trivial but laborious extension of the same reasoning would work if the periods in question were not neighboring. Finally, there can be more than two grades of the resource; the argument says that no higher-cost resource can be used in an optimal program until all lower-cost grades have been exhausted. This conclusion has to be modified, of course, if there are topographical constraints on the pattern of resource use. But within the model, it is safe to take $\theta'(t) > 0$.

I We now return to the main thread, and the model as originally formulated. The problem for society is to find the largest permanently sustainable level of consumption. That means: to find the largest constant consumption with which there can be associated a technologically feasible pattern of resource use and investment over infinite time that will neither use up more than the available stock of resources nor drive the stock of reproducible capital into negative numbers. More tersely: to find the largest constant C such that (1), (2), (4), and (5) have a solution for given K_0 with nonnegative $K(t)$.

This problem can be restated in an equivalent, but more conventional form. Choose a trial constant for C , and find the feasible pattern of resource use and investment that minimizes the cumulative use of resources over all time. Then adjust the value of C until the minimized cumulative resource use is just equal to \bar{R} . More formally, the transformed problem is to

$$\text{minimize } \int_0^{\infty} R(t) dt$$

subject to

$$\dot{K} = K^a R^b - OR - C \tag{6}$$

$$R = f(\theta)\theta, \tag{7}$$

4. Necessary conditions for optimality

with the initial conditions $K(0) = K_0$ and $\theta(0) = 0$, and the tacit requirement that all the unknown time functions be nonnegative. (Remember that $L = 1$.) Necessary conditions for this problem are clearly also necessary for the original consumption-maximization problem. The Lagrangean function, in a form that will make the multipliers $p(t)$ and $q(t)$ nonnegative, is

$$L = R + p(\dot{K} - K^a R^b + \theta R + C) + q[R - f(\theta)\dot{\theta}].$$

The Euler-Lagrange first-order conditions with respect to K , R , and θ are

$$-p/p = aK^{a-1}R^b, \quad (8)$$

$$1 + q = p(bK^a R^{b-1} - \theta), \quad (9)$$

and

$$pR - qf'\theta = d/dt[-qf(\theta)]$$

which, upon carrying out of the differentiation, becomes

$$pR = -f(\theta)\dot{q}. \quad (10)$$

These three equations and the two constraints (6) and (7) are available to determine the five time-functions K , R , θ , p , and q . Four of the five are first-order differential equations, so that the general solution will contain four arbitrary constants. Two of them are used up in making $K(0) = K_0$ and $\theta(0) = 0$; the other two are available to fix the initial values of any two among p , q , and R (the remaining one following from (9) so that the resulting path is actually **optimal**).⁴

As one might have expected, this is a rather more complicated set of first-order conditions than those that arise in the same problem when the resource pool is taken as homogeneous and extraction costs are neglected. That special case can be reproduced by deleting (7) and (10) and setting q and θ both equal to zero in (6) and (9). The point of the present note is to ask what difference it makes to the solution of the resource-management problem to allow for differential extraction costs. To this end we have to interpret the first-order conditions and the dual variables they introduce.

To begin with, p functions as the shadow-price of produced output in terms of the natural resource, so $1/p$ is the shadow-price of resources in terms of **output**.⁵ Then (8) says that, along an optimal path, the shadow-price of resources must increase at a rate equal to the marginal product of reproducible capital in producing output (i.e., itself). This condition is the same whether or not there are differential extraction costs. Its interpretation is direct: $-\dot{p}/p$ is the rate of return from owning and keeping a stockpile of one unit of the resource, in the ground, and the marginal product of capital is the competitive return from owning (and using or renting) reproducible capital, so that along an optimal path an investor must be indifferent between the two forms of holding wealth. One obvious implication is that the shadow-price of the natural resource is always rising.

⁴ These five equations can be reduced to a pair of equations, one second-order and one first-order, in K and θ alone. The initial conditions fix two of the constants of integration and the third has to be chosen, as usual, 'to pick out the trajectory in the (K, θ) plane.

⁵ To be more precise, it is the value of unmined ore.

Equations (9) and (10) do provide something different, however. If there are no extraction costs, $q = \theta = 0$, and (9) simply says that the shadow-price of resources should equal the (shadow-) value of the marginal product of resources in producing output. In the case of differential extraction costs, (9) has to be amended in two ways. First, the relevant marginal product of the resource is its net marginal product, net of extraction costs. That amendment would hold even if all resource deposits were identical; it comes from taking explicit account of the costs of extraction. The second amendment is that the net marginal product of resources must be *larger* than the shadow-price of resources by the factor $(1 + q)$. To put it differently, it is as if an additional cost of q/p units of output per unit of resource consumed were deducted from the net marginal product of resources, and *then* the result set equal to the shadow-price of the resource. Either way, the effect of this amendment is to call for a reduced consumption of resources at any given level of the shadow-price and other things, because there is a shadow-surcharge for the use of the resource over and above its shadow-price.

The interpretation of the surcharge is presumably to be found in (10). Use of (7) to substitute for R leads to

$$\dot{q} = -p\dot{\theta}$$

and, therefore, to

$$q(t) = q(\infty) + \int_t^\infty p\dot{\theta} dt.$$

Evidently, then, q is the shadow-cost of an increase in θ .⁶ A marginal addition to current resource use, apart from using up resources *per se*, uses up the cheapest available resources. At every time in the future θ will be a little larger than it would otherwise have been, and therefore costs a little higher. All those additional costs have to be charged to current resource use and q does that. Dimensionally, q is a pure number measuring cost in resources per unit of resources; q/p is in units of output per unit of resources.

The shadow-cost q is an extension of the “user cost” concept familiar from the literature on exhaustible resources. Bradley (1973), to take a recent example, shows how mere finiteness of a resource implies a shadow-rent to take account of the fact that what is used now will be unavailable for use later. The surcharge q might be called a “degradation cost.” It takes account of the fact that the use of the cheapest available source now implies higher costs for every unit extracted in the future.

■ To get a feel for the implications of extraction-cost differentials, we now examine at length the simplest variation on the model just presented. Suppose that the resource endowment consists of two grades only, a low-cost grade in amount R_1 with cost-parameter θ_1 and a high-cost grade in amount R_2 with cost-parameter θ_2 . It is clear from our earlier analysis that the best plan is to exhaust the low-cost

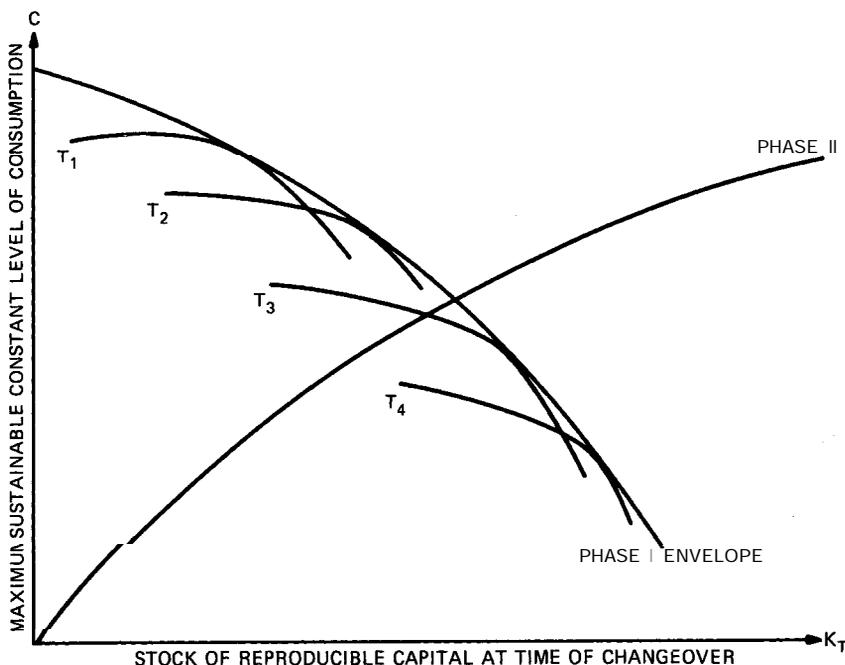
5. A two-grade case

⁶ A necessary transversality condition is that $qf(\theta)$ go to zero as $t \rightarrow \infty$. If $f(\bar{\theta}) > 0$, $q(\infty) = 0$, and perhaps otherwise. In any case, since θ increases in time, q decreases. Geoffrey Heal's closely-related paper (1976) analyzes this very phenomenon under somewhat different technological assumptions.

deposit first, in finite time, and then to spread the high-cost deposit over the infinite remaining future. Call the first part Phase I and the infinite tail Phase II. The switch-over time T must, of course, be chosen according to the criterion of maximizing the permanently-sustainable constant level of consumption per head. We begin by discussing the solution schematically, then give some details, and then some sample computations.

Suppose the low-cost deposit is exhausted at time T : what are the prospects for Phase II? Beginning at time T , the economy has only a homogeneous single grade of the resource left. It must solve an infinite-time maximum-sustainable-consumption problem. (This is exactly the problem solved in Solow (1974), only there it is assumed that $\theta_2 = 0$. We return to this problem later.) The highest level of consumption sustainable in Phase II will obviously be an increasing concave function of K_T , the stock of reproducible capital inherited at time T as a result of the investment decisions made in the first phase of the plan. That function is plotted in Figure 1. Notice that the shape

FIGURE 1
SCHEMATIZATION OF THE TWO-GRADE PROBLEM



of this curve does not depend on T ; only the initial stocks of capital and resources matter, not calendar time. This is the usual dynamic-programming bonus from an infinite horizon.

Now turn to the first phase of the plan. Beginning at time zero, we can imagine the economy solving a homogeneous-resource problem (using only the low-cost deposit) and extending only to an arbitrary time-horizon T , at which time the low-cost deposit is to be exhausted. Given K_0 , the highest level of consumption sustainable during Phase I will be a decreasing function of K_T , the terminal capital stock. There will be one such curve for each T ; in principle, we can draw them all in Figure 1 and their outer envelope will also be a decreasing function of K_T . When K_T is small, the highest achievable level of consumption

will be attained when T is short and production is resource-intensive. If a large K_T has to be accumulated, the highest achievable level of consumption will be attained when T is stretched out to be longer; if the required investment were concentrated in a short interval, even using up the resource quickly would permit only a low level of consumption. (In the extreme case, for any K_T there will be a T so small that even plowing all output back into investment will not succeed in achieving the target.)

Now the intersection of the curve for Phase II with the envelope for Phase I gives the highest permanently-sustainable level of consumption for the combined program, and the stock of reproducible capital to be handed over when society shifts from the low-cost to the high-cost resource. The optimal time of change-over can be read off from the Phase I curve that touches the envelope at the point of intersection. Inspection of Figure 1 is enough to answer some simple questions about the process. For example, a fresh discovery of high-cost resource will rotate the Phase II curve upward and leave the Phase I envelope unchanged; the new optimal program will have higher C , lower K_T , and smaller T . A find of the low-cost resource will shift the Phase I envelope upward, leading to higher C and higher K_T (and, intuition suggests, larger T , but the diagram is silent about that). The computational scheme used later builds on this schematization of the two-grade problem. The details are rather complicated, and are therefore omitted here.

The point of oversimplified models like this is, of course, to give some feel for the character of solutions and their sensitivity to varying assumptions. We have done a few calculations with this goal in mind. The first task is to choose the parameters in a "plausible" way.

In all cases we took $\theta_1 = 0.03$ and $\theta_2 = 0.09$. Thus the low-grade source is three times as costly to mine as the high-grade source. The particular numbers were chosen so that aggregate extraction costs would amount to some 3-5 percent of aggregate consumption when the high-grade source is being mined, proportionately more when the low-grade source is in use. It is hard to get any empirical grip on the bulk of aggregate extraction costs of productive minerals in a modern industrial economy; in retrospect, we think we should also have experimented with some higher values of θ , to see what happens when very high-cost sources have to be called into use.

In all cases we chose $a = 0.2$; we have some experiments with $b = 0.05$, some with $b = 0.1$, and one with $b = 0.15$. If b represents the share of net national product imputed competitively to "land" or natural resources, then 0.05 seems about right. We tried the higher values as well to give some insight into the possible consequences of drastic scarcity of natural resources.

The fixing of initial conditions requires some discussion. Take $t = 0$ to be roughly the present, and write the production function as $Q = AK^aR^b$. Choose units temporarily so that $R = 1$; then current reserves are measured in years of supply at the current rate of extraction, and $Q_0 = AK_0^a$. Now, if the current capital-output ratio is taken as 2, then $K_0^{1-a} = 2A$. If we want to set $A = 1$ for convenience, then we should choose the initial value K_0 equal to $2^{1/(1-a)} = 2^{5/4} = 2.38$. Rounding up, we have set $K_0 = 2.4$ most of the time, with a couple of tries with K_0

6. Some numerical examples

TABLE 1
RESULTS

CASE	C	T	K_T	R_0	R_T
I. $b = 0.05$ $K_0 = 2.4$ $R_1 = 10, R_2 = 50$	1.1405	11	2.76	1.17	0.665
II. $b = 0.05$ $b = 0.10$ $K_0 = 2.4$ $R_1 = 10, R_2 = 50$	1.1385	6.67	2.945	1.825	1.212
III. $b = 0.15$ $K_0 = 2.4$ $R_1 = 10, R_2 = 50$	1.0677	9.8	3.946	1.4238	0.734
IV. $b = 0.05$ $K_0 = 2.4$ $R_1 = 10, R_2 = 25$	1.1390	11.48	2.79	1.14	0.626
V. $b = 0.10$ $K_0 = 2.4$ $R_1 = 10, R_2 = 25$	1.108	9.8	3.30	1.39	0.734
VI. $b = 0.05$ $K_0 = 2.4$ $R_1 = 50, R_2 = 50$	1.1692	36	3.13	1.9175	0.684
VII. $b = 0.05$ $K_0 = 4.8$ $R_1 = 50, R_2 = 50$	1.3035	11.04	5.278	1.075	0.723
VIII. $b = 0.1$ $K_0 = 4.8$ $R_1 = 10, R_2 = 50$	1.2675	8.96	5.76	1.335	0.926

– **4.8**, just to see what doubling the initial capital stock might do. (Perhaps this is a good place to reiterate that we have assumed away population growth and technical progress.)

The only remaining data are the reserves. In most cases, we set $R_1 = 10$ and $R_2 = 50$. Thus, high-grade reserves amount to about ten years' supply at the current rate of use, and low-grade reserves provide another 50 years' worth. It should be realized that the optimal Phase I path does not necessarily start at $R(O) = 1$; in fact in all our examples it starts out using up the low-cost resource faster than that, but then reduces resource-use as capital accumulates. And, of course, the problem is set up precisely so as to stretch out the whole resource endowment over infinite time. We have also experimented with some examples in which R_2 is reduced to 25, and one in which $R_1 = 50$, $R_2 = 50$.

The results are arranged in Table 1. Each "case" is specified by the values of the four parameters that vary from case to case: b , R_1 , R_2 , and K_0 . The successive columns for each case list C (the maximum sustainable constant level of consumption), T (the date of changeover from low-cost to high-cost resource), K_T (the stock of capital at time of changeover), R_0 , and R_T (the rate of use of the resource at the beginning of the optimal program, and at the date of changeover).

The most striking result of these computations is undoubtedly the relative insensitivity of the maintainable consumption level to the size of the resource endowment. Look, for instance, at Cases I, IV, and VI. There b is always set at 0.05, and the initial stock of reproducible capital at 2.4. In I and IV there are 10 units of high-grade reserves; cutting the supply of low-grade resource in half, from 50 to 25,

reduces C by only 0.15 percent. The changeover from low-cost to high-cost reserves occurs only half a year later in Case IV, and the loss of half the stock of high-cost natural resource is made up by the accumulation of an extra one percent of reproducible capital. As between I and VI, Case VI has an endowment of high-grade resource five times as large as that of Case I. But this windfall buys only a 3-percent increase in the permanent level of consumption. The five-fold larger endowment of low-cost resource is stretched out over 36 years in Case VI as compared with 11 years in Case I, and the stock of capital is almost 15 percent larger at the time of changeover.⁷

By contrast, differences in the size of the initial stock of capital make for more substantial differences in the sustainable level of consumption. Case VII is exactly like Case I except that the initial stock of capital is doubled. This allows a level of consumption 15 percent higher. Similarly, Case VIII is like Case II: both have $b = 0.10$, but Case VIII has twice the initial stock of capital. This allows a level of consumption 12 percent higher.

This property of the optimal programs is a consequence of the fact that b , the elasticity of aggregate output with respect to the flow of resource input, is quite small. The flow of resources is a relatively “unimportant” input. Cases I, II, and III are alike, except that b is 0.05, 0.10, and 0.15, respectively. Case II uses up the high-grade resource in less than 7 years, as compared with 11 years in Case I, but accumulates 7 percent more capital in the shorter period. The end result is a trivial reduction in the consumption level. By the time $b = 0.15$, as in Case III, the changeover time is back to 10 years, but in that short period the stock of capital must be increased by 60 percent, compared with 15 percent in Case I and just under 25 percent in Case II. The level of consumption in Case III is 7 percent lower than in I and II, a relatively large difference as Table 1 goes. If b were higher than 0.15, the effect on the maintainable level of consumption would be very dramatic—remember that *if* b were as high as a , namely 0.2, no positive level of consumption would be sustainable for infinite time. An analogous story can be read from a comparison of Case IV with Case V, or Case VII with Case VIII; there b varies only from 0.05 to 0.10.

The importance of this point shows up in a comparison of Cases I and IV on one hand, and Cases II and V on the other. Cases I and IV have $b = 0.05$, but Case IV has its low-grade resource endowment cut in half. As mentioned earlier, there is a trivial reduction in the level of consumption. Cases II and V have $b = 0.10$, and again Case V has its low-grade resource endowment cut in half. This time, by contrast, the maintainable level of consumption falls by 3 percent. That is still small, but it is vastly larger than the minute drop when $b = 0.05$. Unfortunately, we do not have a case to compare with **III**, where $b = 0.15$, but there is no doubt that cutting the low-grade resource endowment in half would then generate a much more substantial penalty in sustainable consumption. So also the contrast between Case VI and Case I would be magnified if b were larger.

⁷ This insensitivity of consumption possibilities to resource endowment probably reflects the fact that our model economy is fairly rich in reproducible capital to begin with. If we were exploring nearer the origin in Figure 1, a change in R_1 (i.e., a shift of the Phase I envelope) would generate a bigger effect on C , by virtue of the concavity of the Phase II curve.

In these infinite-time allocation problems, the significance of exhaustible resources turns in an important way on fine properties of the technology. In this paper, we have confined ourselves to the Cobb-Douglas framework. Within that framework, the output elasticities with respect to resource inputs are the crucial parameters. They are relatively “observable” and casual observation suggests that they are small. From earlier work we know that different elasticities of substitution will generate even more dramatic differences in consumption possibilities. We know, for instance, that if the elasticity of substitution between resource inputs and others is bounded below unity (when the resource input is small), then no positive level of consumption is sustainable forever, whereas if that elasticity of substitution is bounded above unity, then resource flows are not even essential to production. The trouble is that such second-order properties of the technology are not easily “observable” and so we have very little even to guide our intuition. Even more important, at this level of delicacy the aggregative approach becomes insupportable; and so does the assumption that there is no technological progress in production, in extraction, or in the invention of new fuels and materials (and thus, effectively, in the “creation” of new reserves). Nevertheless, a limited guide to our intuition is better than no guide at all.

References

- BRADLEY, P. “Increasing Scarcity: The Case of Energy Resources.” *The American Economic Review*, Vol. 63, No. 2 (May 1973), p. 119.
- HEAL, G. “The Relationship Between Price and Extraction Cost for a Resource with a Backstop Technology.” *The Bell Journal of Economics*, Vol. 7, No. 2 (Autumn 1976).
- SOLOW, R. “Intergenerational Equity and Exhaustible Resources.” *Review of Economic Studies* (Symposium issue) (1974), pp. 29-46.