The Optimal Consumption of Depletable Natural Resources
Author(s): Milton C. Weinstein and Richard J. Zeckhauser
Source: *The Quarterly Journal of Economics*, Vol. 89, No. 3 (Aug., 1975), pp. 371-392
Published by: Oxford University Press
Stable URL: https://www.jstor.org/stable/1885258
Accessed: 27-05-2020 21:10 UTC

THE OPTIMAL CONSUMPTION OF DEPLETABLE NATURAL RESOURCES *

MILTON C. WEINSTEIN
RICHARD J. ZECKHAUSER

I. INTRODUCTION

There is an energy crisis, we are told, facing this nation and the world. It is widely asserted that our economic system leads to an excessive consumption rate for coal, oil, and natural gas, and that something approaching a Malthusian catastrophe may be upon us within our lifetimes or those of our children. This argument is extended elsewhere to relate to wide varieties of extractable minerals and depletable natural resources.[1] An extreme version of this point of view was expressed by ardent conservationist Samuel H. Ordway two decades ago:

> . . . Within foreseeable time increasing consumption of resources can produce scarcities serious enough to destroy our American Dream. . . .[2]

* This research was supported in part by the Analytic Methods Seminar, Kennedy School of Government, Harvard University. We are indebted to Robert Dorfman, Fred Peterson, Howard Raiffa, Thomas Schelling, and a referee for helpful comments.

1. They include, for example, helium and other inert gases. The government's helium conservation program is predicated as a policy response to such arguments.

2. S. Ordway, *Resources and the American Dream* (New York: Ronald Press, 1953), p. 8. H. H. Barnett and C. Morse (*Economics of Resource Scarcity*; Washington, D.C.: Resources for the Future, 1959) provided this quotation and the rest of the historical background material that complements our analysis in this paper. They discuss the following authors: J. Ise ("The Theory of Value as Applied to Natural Resources," *American Economic Review*, XV (1925), 284–91), an early analyst who "finds that the time distribution of the destructive utilization of resources is dangerously biased toward the present"; H. Hotelling ("The Economics of Exhaustible Resources," *Journal of Political Economy*, XXXIX (April 1931, 137–75), who demonstrates the efficiency of the competitive market allocation under assumptions more restrictive than those considered here; and E. O. Heady ("Efficiency in Public Soil Conservation Programs," *Journal of Political Economy*, LIX (Feb. 1951), 47–60), who presents an empirically based analysis of optimal programs for soil conservation.

In response to these fears the government is enacting an increasing number of policy measures that would slow down the rate of consumption of such natural resources. One example of such a measure is the provision in the National Environmental Policy Act that requires all Federal agencies to identify for all proposed projects "any irreversible and irretrievable commitments of resources which would be involved."[3] The clear implication is that the market prices of such commodities do not reflect long-run social cost and that the government must price these commodities de novo in each instance in order to capture their value to the future.[4]

Our purpose is to demonstrate that, under certain standard simplifying assumptions, the consumption stream of a depletable, nonreclaimable, nonreproducible resource that is produced by competitive market behavior does, in fact, coincide with the socially efficient consumption stream. We then extend the analysis to cases where future demand is uncertain and where the costs of extraction may vary. The latter extension enables us to include as well resources that are reclaimable or reproducible, but at prices sufficiently high to make conservation questions of interest. In addition, we examine the behavior of a monopolist seller and demonstrate that in general the direction and magnitude of the departure from optimality cannot be predicted.

The present analysis is directed primarily toward resources whose natural sources can be privately owned. Resources that are publicly owned are handled by the model if a central authority makes appropriate charges for removal of the resource.[5]

II. The Simple Model

We begin with the most restrictive assumptions and then relax a subset of them one at a time, showing how each produces deviations from the initial result. The simple model deals with a resource

3. Public Law 91–190, Section 102 (2) C (V), 1969.

4. Most environmentalists would assert that the government is doing far too little to slow rates of resource depletion. Some policies such as mineral depletion allowances may have the net effect of stepping up resource consumption.

5. In a wide variety of policy relevant cases, insufficient or zero charges are levied for a publicly owned resource. A problem of congestion of the commons is the result. See G. Hardin, "Tragedy of the Commons," *Science,* CLXII (1965), 1243–48. In the natural resource context, this leads to overconsumption. See M. Spence ("A Policy Analysis of International Whaling," in Assorted Fall Term Course Materials: Public Policy 210, Kennedy School of Government, Sept, 1972), for a discussion of how this problem relates to the multination whaling industry.

that is depletable in the strict sense. There exists a fixed quantity Q. Once it is used, it is gone forever. It is nonreproducible.[6] Its costs of extraction and storage for a supplier are zero.

Consumer demand is assumed to be separable by time period. The demand functions are represented as

$$p_t = d_t(q_t), \tag{1}$$

where t is an index over time periods, q_t is the total quantity consumed (not the amount purchased) in period t, and p_t is the price in that period. Consumers have no storage capability. This implies that the between-period cross elasticities of demand are zero. Finally, there is a perfect capital market with a stable, risk-free interest rate r, which reflects the social rate of discount.[7]

The yardstick by which we shall measure the optimality of a consumption stream is the discounted sum of consumer-plus-producer surplus. (To keep units comparable, we must assume that the marginal utility of income is essentially constant.) Represent this sum [8] by

$$S = \sum_{t=0}^{\infty} (1+r)^{-t} \int_0^{q_t} d_t(\xi_t)\, d\xi_t. \tag{2}$$

The optimal consumption stream is therefore given by the sequence $\{q^*_t\}$, which maximizes S, subject to the condition,

$$\sum_{t=0}^{\infty} q_t \leqslant Q. \tag{3}$$

Adjoining (3) to (2) by a multiplier λ and differentiating with respect to each q_t, we get the optimality conditions,

$$d_t(q^*_t) = (1+r)^t \lambda \qquad (t=0,\ 1,\ 2,\ \ldots\ ), \tag{4}$$

which is equivalent to

$$p^*_t = (1+\mathrm{r})^t p^*_0 \qquad (t=0,\ 1,\ 2,\ \ldots\ ), \tag{5}$$

since $\lambda = d_0(q^*_0)$ from substitution of $t=0$ into (4) and since $p^*_t =$

6. This model excludes diamonds because their use does not depreciate or consume them. Heavy metals are excluded because they are reclaimable; timber is excluded because it is reproducible. Of course, if one returns to the atomic level, all resources are reproducible at a price. If the price of reproduction is high enough, the resource fits this simple model. Our initial examples of oil, coal, and natural gas are splendid in this regard.

7. That the market and social rates of discount coincide implies a correspondence between individual and societal valuations of future generations. Society-at-large must have the same trade-off rate as producers and investors between their income and the income of their heirs.

8. The consumer surplus integral will be finite as long as the real income of all individuals is bounded. Even if coal or oil are "necessities," there is a limit to how much people will pay for them. If there are expensive, but feasible, alternative technologies, then the existence of this upper bound becomes more apparent.

$d_t(q^*_t)$ by definition. Thus, (5) specifies that the optimal prices must grow geometrically at the rate r.[9] The initial price is determined by the constraint (3), which, when binding, is written as

$$(6) \qquad \sum_{t=0}^{\infty} d_t^{-1}(p_t) = Q.$$

Note that if the demand curves intersect the vertical axis, then it is possible that q^*_t vanishes after p^*_t rises sufficiently high. In this case it may turn out that the optimal consumption stream is a finite stream, terminating at some period T. This case is consistent with the results of (5) and (6), where $d_t^{-1}(p_t)$ vanishes for $t > T$.[1]

Equation (6) suggests a graphical interpretation of this result that will be useful in describing the market equilibrium conditions. Interpret the left-hand side of (6) as a long-run inverse demand function based on the price at $t=0$,

$$(7) \qquad D^{-1}(p_0) = \sum_{t=0}^{\infty} d_t^{-1}\{p_0(1+r)^t\},$$

where $p_t = p_0(1+r)^t$ has been substituted from (5). Figure I displays the condition that long-run demand must equal long-run supply (the latter being completely inelastic at Q). The optimal initial price p^*_0 is thus determined as the price at which the supply is just exhausted.[2]

9. The analogy between this result and the behavior of von Neumann's model of balanced growth is interesting. In the classical von Neumann model equilibrium prices remain constant over time, while here the price of the fixed-supply resource increases geometrically. The interpretation is that a nonreplenishable resource cannot be part of a "golden age" equilibrium in the von Neumann sense. As the stock of the depletable resource diminishes, it becomes eliminated from the system; the geometric price increase is but a transient phenomenon associated with a transient good. If there are no substitutes for the depletable good, then continued production is impossible, and the only long-run equilibrium will be one of extinction.

1. This situation would occur, for example, if there were substitutes for the good. Suppose, for example, that there exists a perfect substitute that can be produced at a constant marginal cost π. Then the demand curves for the depletable good must intersect the vertical axis at $p_t = \pi$ for all t. The optimal consumption stream in this case would involve utilizing the free but depletable good until it is exhausted (which occurs when its price reaches the price of its substitute π) and then switching to the substitute. If, however, the cost of extracting the depletable resource were nonzero, then it may be optimal to switch partially or totally to the substitute prior to exhaustion of the fixed supply. We see this latter behavior, for example, in the switch from coal to nuclear power.

2. Notice that, in principle, the long-run demand curve in Figure I could intersect the horizontal axis before Q, thus implying that it is optimal *not* to use up all of the resource. This can happen, of course, only if the total demand for the resource at zero price does not add up to Q— an unlikely possibility that we rule out.

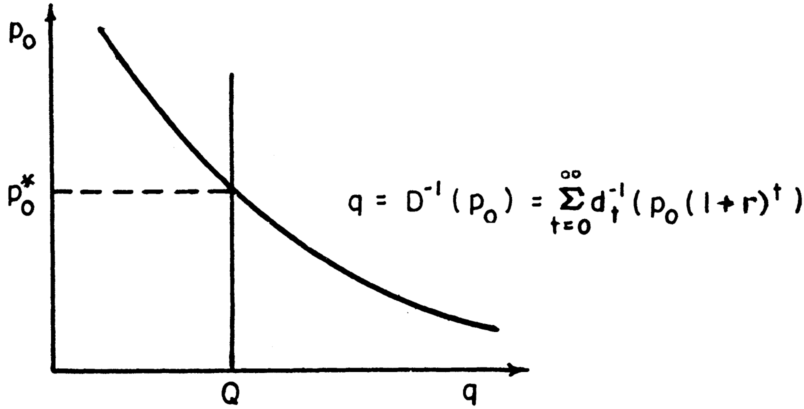

FIGURE I
Long-Run Optimality Condition

III. MARKET EQUILIBRIUM IN THE SIMPLE MODEL

The main result is that, in equilibrium, a competitive market will result in the optimal consumption stream described by (5) and (6). Equivalently, the market price will rise by exactly the factor $(1+r)$ each year, and in the limit the supply will be just exhausted.[3]

Suppose for now that the suppliers have perfect foresight of the consumer demand functions, and further assume that consumers can store the resource at zero cost (so that they could purchase a lifetime supply of the resource now if they chose to do so). These two assumptions will be relaxed presently. The effect of these assumptions is to allow the market to be resolved immediately. On the supply side, consider individual inverse supply functions $S_t^i = f_t^i(Q^i, p_0, p_1, \ldots)$ for the i^{th} supplier in the time period t, where $Q^i = \sum_{t=0}^{\infty} S_t^i$. Since to the supplier a dollar today is always worth $(1+r)$ dollars next period, this function displays infinite cross elasticities along the rays defined by $p_t = p_0(1+r)^t$. In other words, nothing will be marketed in periods when

$$(8) \qquad p_t(1+r)^{-t} < \max_t p_t(1+r)^{-t}$$

3. It should be noted that the optimality and equilibrium results hold even if there are several depletable resources that serve related purposes in the economy (e.g., oil and coal). In particular, it is clear that for any time stream of prices for good A the optimal — and equilibrium — prices for good B will rise at the period rate r. The other good is treated essentially like any other good in the economy. (The strong interdependence between the two will be reflected in starting price levels.) This implies that in equilibrium the prices for each of the two goods will rise at this constant exponential rate.

On the demand side, since it costs nothing to buy in advance, the individual demand functions $D_t^j = h_t^j(p_0, p_1, \ldots)$ will display a similar property. Nothing will be bought in any period t for which there exists an earlier period $t' < t$ such that $p_t > p_{t'}(1+r)^{t-t'}$.[4]

At equilibrium, prices must rise at exactly the rate r. The long-run demand function (7) becomes operational, due to the zero storage costs to the consumers. The equilibrium is determined as in Figure I. The resulting consumption stream is efficient since it satisfies the optimality conditions (5) and (6).

It is more difficult, but nevertheless possible, to see that the same equilibrium holds if either one of the assumptions about perfect information or zero storage cost is not satisfied. If suppliers have perfect foresight of demand, but consumers must use what they purchase in the period of purchase (the most stringent case where their storage costs are effectively infinite), then the individual supply functions remain the same, while the demand functions become the consumption demand functions (1). In this case, the behavior of the suppliers will be sufficient to insure that the price rises at the rate r. Given their perfect foresight of demand, if the quantity supplied in any period were such that the prices could not stay in the fixed geometric growth sequence, suppliers would sell nothing in periods when (8) holds. Such a situation cannot be in equilibrium (unless demand at zero price vanishes in those periods). Thus, the quantities supplied must be such that price rises at the rate r. (This result will be demonstrated more rigorously as a special case of the situation where future demand is uncertain.)

As for the case where producers do not have perfect information, but where consumers can store the commodity at zero cost, we would expect the market to be resolved immediately through the long-run demand function shown in Figure I. Given the free storage, consumers become equivalent to producers in every way, and prices must rise at the rate r in equilibrium. In the case where producers do not have perfect information, *and* where consumers face effectively infinite storage costs, the functioning of a futures market will insure the same equilibrium. However, if future demand is genuinely unknown (even to consumers), then we are in the case of uncertain demand to be treated in Section V.

4. A property of these demand functions is as follows. If prices rise at a rate greater than r, consumers will purchase everything in period zero. If prices rise at a rate smaller than r, however, the actual demand function will coincide with the consumption demand functions (1).

IV. The Extended Model — Optimal and Equilibrium Allocation with Nonzero Extraction Costs

The first major assumption to be relaxed is that the cost of extracting the resource is zero or negligible.[5] We consider first the case where the marginal cost of extraction is constant and then turn to the more realistic case where the marginal cost of extraction increases as the supply diminishes. It will be shown in both cases that the socially optimal consumption stream still coincides with the market equilibrium.

A. Constant Marginal Cost

Let the marginal cost of extraction be constant at m. Then the optimal time stream $\{q^*_t\}$ is the one that maximizes

$$S = \sum_{t=0}^{\infty} (1+r)^{-t} \int_0^{q_t} (d_t(\xi_t) - m)\, d\xi_t, \tag{9}$$

subject to the supply constraint (3). Adjoining (3) to (9) by a multiplier λ, we get the result that

$$d_t(q^*_t) - m = (1+r)^t \lambda \qquad (t = 0, 1, 2, \ldots),$$

which is equivalent to

$$p^*_t - m = (1+r)^t (p^*_0 - m) \qquad (t = 0, 1, 2, \ldots). \tag{10}$$

Thus, (10), in a manner analogous to (5), specifies that price minus marginal cost should grow at the rate r. Note that this implies that price itself grows more slowly, but that the rate of price rise approaches r in the limit.

It is clear that this is exactly what happens in equilibrium. Suppliers are indifferent between a dollar of profit now and $1+r$ dollars in the next period, so that the amount supplied will be nonzero only if price minus marginal cost grows at the rate r. On the demand side it is irrelevant whether the commodity can be stored. Since the price rises more *slowly* than r, consumers will purchase the amount they wish to consume in each period. The long-run inverse demand curve is then given by

$$D^{-1}(p_0) = \sum_{t=0}^{\infty} d_t^{-1}(p_t), \tag{11}$$

5. It is worth noting here that there is an analytical analogy between this extended model and another model that would take into account the possibility of recycling, where the cost of recycling plays the role of the cost of extraction. The relationship between market outcomes and optimal allocations when there is a possibility for recycling is discussed in M. Weinstein and R. Zeckhauser, "Use Patterns for Depletable and Recycleable Resources," *Review of Economic Studies, Symposium*, XLII (1974), 67–88.

where p_t satisfies $(p_t - m) = (p_0 - m)(1+r)^t$. The equilibrium is shown in Figure II. The equilibrium initial price p^*_0 is determined

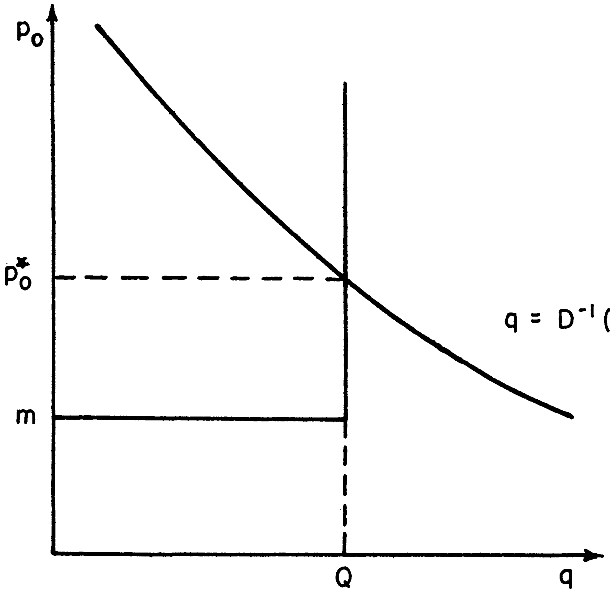

FIGURE II
Long-Run Optimality Condition with Constant Marginal Extraction Cost

by (6) with p_t appropriately redefined, so that $p_t - m$, not p_t, grows at the rate r. Note that it is possible that the long-run demand curve may cut the marginal cost curve at some $q' < Q$ as in Figure III. In this case the initial price is set equal to marginal cost, and through (10) this implies that the price always remains at marginal cost. What is happening here is that the supply constraint is not binding; the shadow price $\lambda = p_0 - m$ is zero. Intuitively, this reflects a situation where the cost of extraction runs ahead of demand so that it does not pay to use the total supply. In reality such a situation may exist for such resources as sand and gravel, and possibly coal, but it is unlikely for oil and natural gas. Indeed, for resources like sand, the cost of extraction and delivery may be so great that the effective value of the resource in its natural state may be essentially zero.[6]

6. Consider, for example, the value of an acre of sand under the ocean or of an acre of unexplored wilderness during the Gold Rush.

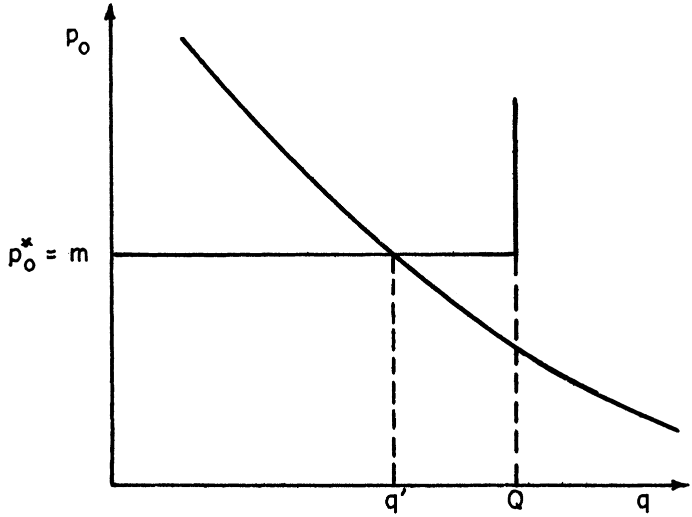

FIGURE III
Incomplete Utilization of the Resource with Constant Marginal Cost

B. Increasing Marginal Cost

Now suppose that the *total* cost of extracting the first q units is given by $c(q)$, regardless of when the extraction takes place.[7] The marginal cost of removing a unit is $c'(q)$ for $0 \leqq q \leqq Q$. Each producer has his own marginal cost curve $c'_k(q^k)$ that depends only on the amount q^k that he has extracted to date. The aggregate marginal cost curve is derived as the horizontal sum of these individual curves.

The socially optimal sequence $\{q_t\}$ is the one that maximizes

$$S = \sum_{t=0}^{\infty} (1+r)^{-t} \left[\int_0^{q_t} d_t(\xi_t)\, d\xi_t - c\left(\sum_{i=0}^{t} q_i \right) + c\left(\sum_{i=0}^{t-1} q_i \right) \right],$$

subject to the supply constraint (3). Adjoining the constraint and differentiating, we get the optimality conditions,

$$(12) \qquad d_t(q^*_t) - c'\left(\sum_{i=0}^{t} q^*_i \right) - \sum_{j=1}^{\infty} (1+r)^{-j} \left[c'\left(\sum_{i=0}^{j} q^*_{i+t} \right) - c'\left(\sum_{i=0}^{j-1} q^*_{i+t} \right) \right] = (1+r)^t \lambda,$$

where, by substitution of $t=0$,

7. Note that under these assumptions, it is always optimal for suppliers with perfect information to postpone extraction until the period of sale.

(12a) $$\lambda = d_0(q^*_0) - c'(q^*_0) - \sum_{j=1}^{\infty} (1+r)^{-j}[c'(\sum_{i=0}^{j} q^*_i) - c'(\sum_{i=0}^{j-1} q^*_i)].$$

The interpretation of the left-hand side of (12) is as follows. The first term $d_t(q^*_t) = p^*_t$ is the price. The remaining terms represent the discounted total of increased costs entailed by supplying an additional unit in period t. The second term gives the within-period marginal cost, as it is conventionally understood, of supplying this unit. But given that the cost-of-extraction curve is rising, supplying this unit will increase all future extraction costs as well. Thus, the third term gives the discounted sum of these increments to cost in all future periods. The two terms together give the true, for all time, marginal cost of supplying an additional unit in this period.[8] Thus, (12) and (12a) express essentially the same condition we saw in (10): price minus marginal cost grows at the rate r.

The market equilibrium yields the same allocation as the social optimum. Suppliers will set price minus "marginal cost" equal to the shadow price λ, and the market will clear only when the initial price p^*_0 is set so that

$$D^{-1}(p^*_0) = \sum_{t=0}^{\infty} d_t^{-1}(p^*_t) = Q.$$

Again it is possible that behavior analogous to that shown in Figure III may occur, in which case $\lambda = 0$, and price equals marginal cost (in the extended sense defined by (12)) both in equilibrium and at the optimum.

The efficiency of the market allocation can be seen as follows. Recall that the aggregate marginal cost function $c'(q)$ is the *horizontal* sum of the individual marginal cost functions $c'_k(q^k)$. Now the individual supplier faced with a sequence of price $\{p_t\}$ and a fixed supply Q^k will clearly allocate his resources to satisfy

(13) $$p_t - c'_k(\sum_{i=0}^{t} q_i^k) - \sum_{j=1}^{\infty} (1+r)^{-j}[c'_k(\sum_{i=0}^{j} q^k_{i+t}) - c'_k(\sum_{i=0}^{j-1} q^k_{i+t})] = (1+r)^t\lambda.$$

8. Most discussions of marginal cost need not be concerned with increments to future costs because they assume that marginal cost curves in individual periods are independent of one another.

Parking fine structures in many cities reflect the interdependent cost structure of the example in the text, at least over a one-year period. The first violation costs zero dollars, the next is five, then ten, and so on. A friend borrows your car at the beginning of the year and secures you your first ticket. Clearly, although his violation carries no fine, he is imposing on you an increased cost for all future violations (until your slate is wiped clean). If his compensation to you is to "make you whole," it must equal the discounted total of these increments to fines.

Note that this equation is identical to (12) with c' replaced by c'_k, and the q^*_i replaced by q_i^k. But since c' is merely the horizontal sum of the c'_k,

$$c'(\sum_{i=0}^{t} q_i) = c'(\sum_{i=0}^{t} \sum_{k} q_i^k) = c'_k(\sum_{i=0}^{t} q_i^k),$$

where $q_i = \sum_k q_i^k$ by definition. Thus, the market equilibrium condition (13) reduces to the optimizing conditions (12), and the allocation is seen to be efficient.[9]

V. Equilibrium with Uncertain Future Demand

Let us return to the situation where extraction costs are zero, but now suppose that the demand functions in future periods are unknown and are revealed at the start of each period. We now examine the properties of the equilibrium consumption stream (or the equilibrium probability distribution on consumption streams) when future demand is uncertain. We begin with a two-period model, which illustrates some of the characteristics of the general result, and then turn to a three-period model, which generalizes easily to N periods. The generalization to infinite horizon is not given here.

A. Two-Period Model

Let Q^i be the total quantity owned by the i^{th} producer. Let Q_t be the total quantity supplied in period t, and let q_{it} be the quantity supplied in period t by the i^{th} producer. Let the demand functions in the two periods be given by

$$p_0 = d_0(Q_0)$$

and

$$p_1 = d_1(Q_1, z_1),$$

where z_1, is a random variable of arbitrary dimension. Suppose that the i^{th} producer has a von Neumann-Morgenstern utility func-

9. In an interesting but anomalous second model of production, each supplier owns his own stock of resource, but they all collectively face a joint cost function $c(q)$, which measures the cost of extracting the first q units, regardless of which supplier extracts them. Under these circumstances, we might expect that suppliers would prematurely extract the resource, rushing to pass the "externality" of higher future extraction costs on to others. In fact, this does not happen. The market equilibrium satisfies (12), even though (13) no longer applies. The intuitive reason for this surprising result is that the "externality" of passing higher costs on to other suppliers is not a physical or technological externality for the industry as a whole; the real cost of extracting the first q units remains $c(q)$ no matter how extraction occurs. Since there is no real externality, it is not surprising after all to find that the market allocates the resource efficiently.

tion u_i on discounted assets. Then the i^{th} producer will select q_{i0} to maximize $Eu_i(p_0 q_{i0} + \boldsymbol{p}_1(1+r)^{-1}q_{i1})$ subject to $q_{i0}+q_{i1}=Q^i$. Since $\boldsymbol{p}_1$ depends only on $\boldsymbol{z}_1$ and on $Q_1=Q-Q_0$, this maximization defines a function f_i such that

(14) $\quad q_{i0}=f_i(p_0,Q_0)$, all i.

Market equilibrium is determined by (14) together with the conditions,

(15) $\quad Q_0 = \sum_i q_{i0}$

and

(16) $\quad p_0=d_0(Q_0)$.

If there are I producers, then (14)–(16) form $I+2$ equations in the $I+2$ unknowns $\{q_{i0}\}$, Q_0, and p_0.

Consider now some special cases. If $\boldsymbol{p}_1$ is known (i.e., if $\boldsymbol{z}_1$ has zero variance), then the supplier's maximization yields a function f_i such that

$$q_{i0}=\begin{cases} Q^i & \text{if } p_1<p_0(1+r) \\ 0 & \text{if } p_1>p_0(1+r) \\ \text{indeterminate} & \text{if } p_1=p_0(1+r). \end{cases}$$

In order for demand in both periods to be met, (i.e., in order for a nonzero quantity to be supplied in both periods), it must be the case that $p_1=p_0(1+r)$. This was the result presented for the certainty case in Section III, and this is the efficient allocation.

Suppose now that $\boldsymbol{z}_1$ is unknown but has probability distribution that all producers agree upon. Suppose further that producers are risk-neutral. Then producers will choose q_{i0} to maximize

$$\begin{aligned} J &= E(p_0 q_{i0}+\boldsymbol{p}_1(1+r)^{-1}q_{i1}) \\ &= p_0 q_{i0} + (E\boldsymbol{p}_1)(1+r)^{-1}q_{i1}. \end{aligned}$$

Suppliers will therefore select

$$q_{i0}=\begin{cases} Q^i & \text{if } E\boldsymbol{p}_1<p_0(1+r) \\ 0 & \text{if } E\boldsymbol{p}_1>p_0(1+r) \\ \text{indeterminate} & \text{if } E\boldsymbol{p}_1=p_0(1+r), \end{cases}$$

so that at equilibrium, we must have

(17) $\quad E\boldsymbol{p}_1=p_0(1+r)$.

In other words, if suppliers are risk-neutral, the initial ($t=0$) allocation is such that the *expected* price rises at the rate r.

It is easy to show that if society wishes to maximize expected discounted surplus,

(18) $\quad E[S]=E[\int_0^{q_0} d_0(\xi_0)d\xi_0 + \int_0^{q_1} d_1(\xi_1,\, \boldsymbol{z}_1)d\xi_1],$

then the optimal allocation is such that (17) holds, where

$$p_1 = d_1(q_1, \theta_1)$$

and

$$p_0 = d_0(q_0).$$

Thus, if suppliers are risk-neutral, the market equilibrium is optimal in terms of $E[\boldsymbol{S}]$. This is demonstrated more generally for the multi-period case in the Appendix.

Suppose finally that suppliers are risk-averse. In this case the market equilibrium will result in an expected price rise by a factor *larger* than $(1+r)$. Thus, risk-averse suppliers will underconserve the resource, relative to the numeraire $E[\boldsymbol{S}]$.[1] To see this, consider the supplier who selects q_{i0} to maximize

$$J = Eu_i(p_0 q_{i0} + \boldsymbol{p}_1(1+r)^{-1} q_{i1}),$$

subject to $q_{i0} + q_{i1} = Q^i$. Suppose that

$$E(\boldsymbol{p}_1) = p_1^{\mu} \text{ and Var } (\boldsymbol{p}_1) = p_1^{\nu}.$$

Now if $p_1^{\mu} = p_0(1+r)$ exactly, then the supplier will set $q_{i0} = Q^i$ and $q_{i1} = 0$ because he can never achieve a higher expected value than $p_0 Q^i$ but can minimize his variance by selling all of his supply at the certain price p_0. This situation cannot be in equilibrium, since nothing would be left for period 1. In equilibrium we must have

$$p_1^{\mu} p_0(1+r) \tag{19}$$

in order for suppliers to save any of their stock for the future.

As an illustration, suppose a supplier has a constant risk aversion utility function for discounted revenue r given by

$$u_i(r) = -e^{-c_i r}.$$

Suppose further that $\boldsymbol{p}_1$ is distributed normally with mean p_1^{μ} and variance p_1^{ν}. Then, if the supplier allocates q_{i0} to period 0 and q_{i1} to period 1, his certainty equivalent[2] is given by

$$CE = c_i(c_i p_0 q_{i0} + p_1^{\mu}(1+r)^{-1}(Q^i - q_{i0}) - c_i p_1^{\nu}(Q^i - q_{i0})^2/2(1+r)^2).$$

1. Strictly speaking, it is not appropriate to put a cardinal utility function on unknown future streams of *increments* to consumption (i.e., incomes), as distinguished from streams of pure consumption. This is because decisions concerning interperiod transfers (i.e., borrowing and lending) must occur *before* resolution of uncertainty in the future. The proper evaluation of uncertain income streams requires the solution of a complex dynamic programming problem in which future decisions are considered. See M. Spence and R. Zeckhauser, "The Effect of the Timing of Consumption Decisions and the Resolution of Lotteries on the Choice of Lotteries," *Econometrica*, XL (March 1972), 401–03. Nevertheless, it is a common and convenient practice to use utility of discounted income as a surrogate for utility for consumption.

2. Using moment-generating functions, it is straightforward to show that a gamble with mean μ and variance σ^2 has a certainty equivalent $c\mu - c^2\sigma^2/2$, where c is the parameter of the exponential utility function (i.e., the risk aversion).

This is maximized when [3]

$$Q^i - q_{i0} = (p_1^{\mu} - p_0(1+r))(1+r)/cp_1^{\nu}.$$

Thus, if $p_1^{\mu} = p_0(1+r)$, then $q_{i0} = Q^i$, and all is sold in period O. In equilibrium, the (p_1^{μ}, p_1^{ν}) pair must be such that the market just clears in period 0. This can happen only if $p_1^{\mu} > p_0(1+r)$. Note that, unlike the case where suppliers are risk-neutral, there is no clear-cut value of p_1^{μ} below which all is sold in the first period and above which all is sold in the second period. Here, there is a range of values of p_1^{μ} (for given p_1^{ν}) for which nonzero quantities are supplied in both periods. There is generally, however, a unique (p_1^{μ}, p_1^{ν}) combination at which the two-period market equilibrates.

In summary, the results of the two-period market under uncertainty are the following. If suppliers are risk-neutral, then expected price rises by a factor of $(1+r)$, which is optimal in terms of expected discounted surplus. If suppliers are risk-averse, then expected price rises by a factor greater than $(1+r)$, which results in underconservation of the resource relative to the social optimum in terms of expected discounted surplus.

B. Multiperiod Model

In extending the results obtained in the two-period case to three periods and more, we encounter a qualitative difference in the decision making that underlies the market equilibrium. Where demand is uncertain at least two periods into the future, fully rational suppliers will be forced to turn to closed-loop, or adaptive, dynamic programming, rather than simple open-loop optimization where allocations in all periods would be determined at the start.[4]

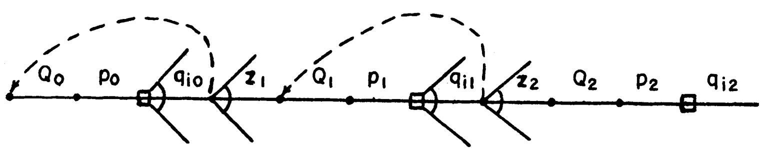

FIGURE IV
Three-Period Decision Tree

3. We ignore boundary solutions, since the market equilibrium must generally involve nonzero sales in both periods.

4. In the two-period case, of course, the two methods are identical because there is only one free decision variable. See R. E. Bellman and S. C. Dreyfus, *Applied Dynamic Programming* (Princeton, N.J.: Princeton University Press, 1962).

Each supplier faces a decision tree as in Figure IV. We assume that the probability distributions for the $z_i (i=1, 2)$ are independent,

The sequence of events is as follows. Given full information about the market in period 0, the i^{th} supplier chooses to sell q_{i0} at the market price p_0. After all suppliers have chosen their $\{q_{i0}\}$ values, the market will be in equilibrium only if $Q_0 = \sum_i q_{i0}$ (indicated by the dashed line in Figure IV). In the next period the uncertainty about demand in period 1 (z_1) is resolved. Then the process repeats itself, and equilibrium is reached in period 1. Finally, the uncertainty about period 2 (z_2) is resolved, and the remaining events are predetermined since the rest of the supply $(Q-Q_0-Q_1)$ must be sold at the revealed market price p_2.

The above producer-optimization *cum* equilibrium system results in a probability distribution over the space of possible allocations of the supply Q. Fortunately, it is possible to say a great deal about the resulting distribution. These results are now demonstrated in the three-period case, only because the notation gets out of hand as the number of periods grows. The same results hold in the N-period case.

Consider the situation at the start of period 1 (i.e., the second period) after z_1 is revealed. *Given* z_1 (and Q_0), the resulting equilibrium consisting of $\{q_{i1}\}$, Q_1, and p_1 is known (since it is no more than the resultant of a two-period process starting with supply $Q-Q_0 = \sum_i (Q^i - q_{i0})$). Thus, the *functions,*

$$p_1(z_1, Q_0),\ Q_1(z_1, Q_0),\ \text{and}\ \{q_{i1}(z_1, q_{i0})\}, \tag{20}$$

are known at the start of period 0. Furthermore, the probability distribution for

$$p_2 = d_2(z_2, Q_2) = d_2(z_2, Q-Q_0-Q_1[z_1, Q_0]) \tag{21}$$

is also known at the outset by (20). The individual supplier will choose his initial q_{i0} to maximize

$$\begin{aligned} J = E_{z_1,z_2} u_i [& p_0 q_{i0} + (1+r)^{-1} p_1[z_1, Q_0]\, q_{i1}[z_1, q_{i0}] \\ & + (1+r)^{-2} d_2, [z_2, Q-Q_0-Q_1[z_1, Q_0]]\,(Q^i - q_{i0} - q_{i1}[z_1, q_{i0}])], \end{aligned} \tag{22}$$

where the dependence of all prices and quantities on z_1 and z_2 is shown explicitly. The expression (22) is a function of q_{i0}, Q_0, and p_0, so that the optimal q_{i0} is a function of Q_0 and p_0:

$$q_{i0} = f_i(Q_0, p_0).$$

The equilibrium is completed by the conditions,

$$Q_0 = \sum_i q_{i0}$$

and

$$p_0 = d_0(Q_0).$$

Consider the properties of this equilibrium in some important special cases. In the case where θ_1 and θ_2 are *known*, the maximand (22) reduces to

$$J=p_0q_{i0}+(1+r)^{-1}d_1[Q_0]q_{i1} \\ +(1+r)^{-2}d_2[Q-Q_0-Q_1](Q^i-q_{i0}-q_{i1}).$$

We know in advance, however, that $p_2=p_1(1+r)$, so this reduces further to

$$J=p_0q_{i0}+(1+r)^{-1}d_1[Q_0](Q^i-q_{i0}).$$

As in the two-period case, the supplier will choose

$$q_{i0}=\begin{cases} Q^i & \text{if } p_1<p_0(1+r) \\ 0 & \text{if } p_1>p_0(1+r) \\ \text{indeterminate} & \text{if } p_1=p_0(1+r), \end{cases}$$

and market equilibrium holds, in general, only if $p_1=p_0(1+r)$. Therefore, as claimed in Section III, the market equilibrium yields

$$p_i=p_0(1+r)^i$$

in the certainty case.

If the z_i are uncertain, but suppliers are risk-neutral, then it is an easy matter to deduce the result that the *expected* price rises at the rate r; that is,

$$E_{z_1,z_2}\boldsymbol{p}_2=(1+r)E_{z_1}\boldsymbol{p}_1=(1+r)^2p_0. \tag{23}$$

To see this, note that the two-period equilibrium at the start of period 1 must satisfy the left-hand equality in (23). Therefore, by "folding back" one more step using (22), it is clear that either all or none will be sold in period 0 unless the right-hand equality in (23) also holds. This result extends easily to N periods, the general result being that

$$E\boldsymbol{p}_i=(1+r)^ip_0. \tag{24}$$

It turns out that this market equilibrium is socially optimal in terms of expected discounted surplus,

$$E[\boldsymbol{S}]=E[\int_0^{Q_0}d_0(\xi_0)d\xi_0+\sum_{t=1}^{\infty}(1+r)^{-t}\int_0^{Q_t}d_t(\xi_t,\ z_t)d\xi_t], \tag{25}$$

where society solves its own closed-loop (adaptive) dynamic programming problem. This result is given in the Appendix.

Finally, it is worth noting that, as in the two-period case, if suppliers are risk-averse, the expected price will rise at a (not necessarily constant) rate *higher* than r, so that, judged by the numeraire $E[\boldsymbol{S}]$ in (25), the resource is being used up too quickly.

VI. Monopoly Behavior

When Progressivism and Conservationism were flourishing side by side in the early part of this century, it was generally believed that monopolies were the enemies of conservation. Monopolists, the common wisdom went, would exploit our natural resources for their own profits at a rate too fast for the good of society.

A closer examination of the economics of monopoly, however, leads one to suspect that the tendency of a monopolist to restrict supply below optimal levels would mean that the resource, if monopolized, would actually be overconserved.[5] If the truth be known, depending upon circumstances a monopolist may underconserve, overconserve, or optimally conserve a resource. For expected sets of circumstances, however, the tendency is toward overconservation.

The simple model illustrates the possibilities. Extraction costs are zero; there is perfect foresight of demand. The monopolist must choose the sequence $\{Q_t\}$ to maximize revenue,

$$(26) \qquad R = \sum_{t=0}^{\infty} (1+r)^{-t} d_t(Q_t) Q_t,$$

subject to the constraint,

$$(27) \qquad Q \geqq \sum_{t=0}^{\infty} Q_t.$$

Adjoining (27) to (26) by a multiplier λ and differentiating with respect to the $\{Q_t\}$, we see that the monopolist's constrained optimum occurs when

$$(28) \qquad d_t(Q_t) + d'_t(Q_t) Q_t = (1+r)^t \lambda.$$

Recognizing the left-hand side of (28) as marginal revenue in period t (MR_t) and solving for the initial condition at $t=0$, we get the condition,

$$(29) \qquad MR_t = (1+r)^t MR_0,$$

so that *marginal revenue*, not price, grows at the rate r. If the constraint (27) is not binding (an unlikely occurrence in reality), then $\lambda = 0$, and the monopolist will sell in each period up to the point where marginal revenue is zero.[6]

5. T. Schelling ("Monopolistic Restriction and the Production of Bads," mimeograph, Kennedy School of Government, Harvard University, 1972) makes a similar observation in suggesting that a monopolistic industry may produce fewer public bads (i.e., pollution) than it would if it were competitively organized.

6. If this occurred, then at least in the long run the monopolist would overconserve, since it is never socially optimal to fail to meet (27) with equality.

Growth of marginal revenue at rate r implies nothing about social optimality. The key factor is the growth rate of price. Consider cases where the monopolist exhausts his stock $(\lambda > 0)$. If price grows less swiftly than r, he is overconserving. If it grows at the rate r, his allocation is socially optimal. If price rises faster than r, he is underconserving. In cases where he does not exhaust his stock $(\lambda = 0)$, the result is ambiguous if price rises faster than r, otherwise he is definitely overconserving.

Suppose first that the demand curve in period t is given by

$$p_t = a_t - b_t Q_t,$$

so that the marginal revenue is given by

$$m_t = a_t - 2b_t Q_t.$$

Hence,

$$p_t = (a_t + m_t)/2,$$

independent of the slope b_t. Suppose that m_t is growing at the rate r per period, according to the monopolist's optimum. Then

$$p_t(1+r) - p_{t+1} = \frac{a_t(1+r) + m_t(1+r)}{2} - \frac{a_{t+1} + m_t(1+r)}{2}$$
$$= \frac{a_t(1+r) - a_{t+1}}{2}.$$

Therefore, if the *intercept* a_t is growing at a rate less than r, then price also grows at a rate less than r, and the resource is overconserved. Otherwise, both society and the monopolist would seek to preserve the resource from consumption indefinitely. In practice, it seems unlikely that the intercept (i.e., the price above which none of the good is demanded) will grow as rapidly as the interest rate. It is more likely that demand would shift horizontally by a factor representing population growth, this having no effect on the intercept. Therefore, in the linear case we are inclined to conclude that a monopolist will *overconserve* the resource.

Now consider the case of constant elasticity demand:

$$p_t = A_t q_t^{-\alpha_t} \qquad (0 < \alpha_t < 1).$$

In this case, marginal revenue is proportional to price:

$$m_t = (1 - \alpha_t) p_t.$$

Now if the inverse elasticity α_t is constant over time, then a geometric increase in m_t at the rate r implies a geometric increase in price at the same rate. Therefore, with constant and stable elasticity of demand, the monopolist's allocation coincides with the market equilibrium and is socially optimal.[7] If the inverse elasticities a_t increase (decrease) (remain constant) with time, then price increases at a rate greater than (less than) (equal to) r if marginal revenue increases at the rate r. In this case the monopolist underconserves (overconserves) (conserves optimally).

Only empirical study can determine what set of assumptions about demand for a monopolized resource is satisfied in a particular instance. Once that is known, as these examples illustrate, it can be determined whether a monopolist is guilty of overprotecting or underprotecting the resource under his control. It is just possible that he is behaving optimally in this respect.

VII. Summary and Conclusion

It has been shown that a perfectly competitive market for a depletable natural resource will, under certain conditions, result in efficient intertemporal allocation. This allocation pattern is characterized by an exponential price increase at the marginal rate of time preference in the society. Where there is a positive extraction cost, exponential growth will be in the difference between price and the marginal extraction cost, appropriately measured.

The conditions specified include that the participants have access to perfect capital markets, that the resource can be privately owned, and that there are no unpriced externalities. If these conditions are not satisfied, there may be an argument for government participation, for example, to encourage mineral production that provides an externality for national security, to license a fishing area that would otherwise be overharvested, or to impose a variety of conservation measures because an excess in suppliers' effective interest rates over the marginal rate of time preference leads them to deplete the resource too rapidly.

Uncertainty about future demand has no adverse effect on optimality if suppliers are risk-neutral. If suppliers are risk-averse, but society is risk-neutral, then the resource will tend to be underconserved. The presence of monopoly may produce either under- or overconservation, though overconservation would be the expected result.

In general, despite the presence of factors that prevent us from

7. Provided that the supply constraint (27) is binding. We would expect the constraint to be binding in general, since if it were not then the monopolist would sell up to the point where marginal revenue is zero in each period and never use up his supply. In the constant elasticity case this is impossible since marginal revenue never reaches zero.

fully generalizing our efficiency findings, we feel that it is appropriate to conclude on a note of cautious optimism and to point out that some recent doomsday predictions about our inevitable bare cupboards seem overdrawn.[8] It is powerful solace to know that underlying market forces will work to produce appropriate rates of resource consumption. Fisher and Potter conclude that

> There will, indeed, be supply problems for particular resources at particular times and places; but technological and economic progress, building upon an ample and diversified resource and industrial base, gives assurance that supply problems can be met.[9]

Finally, we can think of no more appropriate oracle than John von Neumann himself, who, confronted with the dawn of the nuclear age, wrote that

> It is likely that we shall gradually develop procedures more naturally and effectively adjusted to the new sources of energy, abandoning the conventional kinks and detours inherited from chemical-fuel processes. Consequently, a few decades hence energy may be free — just like the unmetered air — with coal and oil used mainly as raw materials for organic chemical synthesis, to which, as experience has shown, their properties are best suited.[1]

Appendix: Proof of the Optimality of the N-Period Market Equilibrium Under Uncertainty If Suppliers Are Risk-Neutral

It was shown in Section VI that the market equilibrium in the case where future demand is uncertain but where suppliers are risk-neutral is characterized by the property that the *expected* price rises at the interest rate r:

(i) $$p_0 = (1+r)^{-t} E\boldsymbol{p}_t.$$

It was claimed then, and is proven here, that such an allocation strategy is optimal from the point of view of maximizing expected discounted surplus:

(ii) $$E[S] = E\Big[\int_0^{Q_0} d_0(\xi_0)\, d\xi_0 + \sum_{t=1}^{\infty} (1+r)^{-t} \int_0^{Q_t} d_t(\xi_t, z_t)\, d\xi_t\Big],$$

8. A good gloomy example is provided by D. Meadows *et al., The Limits of Growth: A Report of the Club of Rome's Project on the Predicament of Mankind* (New York: Universe, 1972).
9. J. L. Fisher and N. Potter, "The Effects of Population Growth on Resource Adequacy and Quality," in *Rapid Population Growth: Consequences and Policy Implications*, Vol. 2 (Baltimore: National Academy of Sciences, 1971), p. 224.
1. J. von Neumann, "Can We Survive Technology?" in *The Fabulous Future*, (New York: E. P. Dutton and Co., 1955), p. 37.

subject, of course, to the constraint that

(iii) $$Q = \sum_{t=0}^{N} Q_t.$$

The proof given here is by induction. The result is first proved in the case $N=1$ and then generalized to $N=K$ by assuming the result to be true for $N=K-1$. It will be useful for notational purposes to suppress the expectation operator and to express the integration over the parameter spaces θ_t explicitly.

The two-period proof is straightforward, since there is only one decision variable Q_0 and no possibility for adaptive control. We want to choose Q_0 to maximize

$$E[S] = \int_0^{Q_0} d_0(\xi_0)\,d\xi_0 + (1+r)^{-1} \int_{z_1} \int_0^{Q-Q_0} d_1(\xi_1, z_1)\,d\xi_1 dz_1.$$

Taking the derivative with respect to Q_0 and setting it equal to zero yields

(iv) $$d_0(Q_0) - (1+r)^{-1} \int_{z_1} d_1(Q_1, z_1)\,dz_1 = 0,$$

where $Q_1 = Q - Q_0$. Rewriting (iv) in terms of expected values and prices, we see that

(v) $$p_0 = (1+r)^{-1} E[\boldsymbol{p}_1],$$

which completes the proof for the case $N=1$.

Now suppose that the result is true for $N=K-1$. Then the induction hypothesis may be written in terms of the *last* K periods of the case $N=K$ as follows:

(vi) $$p_1 = (1+r)^{1-t} E(\boldsymbol{p}_t) \qquad (t = 1, \ldots, K).$$

We wish to show that this extends backwards to period 0, i.e., that

(vii) $$p_0 = (1+r)^{-t} E(\boldsymbol{p}_t) \qquad (t = 1, \ldots, K).$$

Before embarking on this proof, it must be noted that in this closed-loop optimization, only Q_0 must be chosen. The remaining Q_t may be chosen in the future, conditionally on information received up to that time.[2] Thus, Q_t is actually a *function* of $Q_0, \ldots, Q_{t-1}$ and $z_1, \ldots, z_t$:

(viii) $$Q_t = Q_t(Q_0, \ldots, Q_{t-1}; z_1, \ldots, z_t).$$

In the maximization procedure it will be necessary to differentiate Q_t with respect to Q_0. We shall use the notation dQ_t/dQ_0 to denote the *total* derivative of (viii) with respect to Q_0 (including the dependency through $Q_1, \ldots, Q_{t-1}$) and not just the partial derivative with respect to the first argument. Thus for example,

$$\frac{dQ_2}{dQ_0} = \frac{\partial Q_2}{\partial Q_0} + \frac{\partial Q_2}{\partial Q_1}\frac{\partial Q_1}{\partial Q_0}.$$

2. The optimal open-loop allocation (which is suboptimal in general) will generally be different from the optimal closed-loop solution. In general, the open-loop solution will be characterized by too much early consumption and not enough conservation. It is interesting to note that although the open-loop solution is still characterized by expected price rises at the rate r, the solutions do differ.

We are now ready to proceed. The maximand is

(ix)

$$E[S] = \int_0^{Q_0} d_0(\xi_0)\,d\xi_0 + (1+r)^{-1} \int_{z_1} \int_0^{Q_1} d_1(\xi_1,z_1)\,d\xi_1 dz_1 + \ldots$$
$$+ (1+r)^{1-K} \int_{z_1} \ldots \int_{z_{K-1}} \int_0^{Q_{K-1}} d_{K-1}(\xi_{K-1},z_{K-1})\,d\xi_{K-1} dz_{K-1} \ldots dz_1$$
$$+ (1+r)^{-K} \int_{z_1} \ldots \int_{z_K} \int_0^{Q-Q_0-\ldots-Q_{K-1}} d_K(\xi_K,z_K)\,d\xi_K dz_K \ldots dz_1,$$

where the Q_t are actually functions as in (viii). Differentiating (ix) completely with respect to Q_0 and setting the derivative equal to zero, we get

(x)

$$d_0(Q_0) + (1+r)^{-1} \int_{z_1} d_1(Q_1,z_1)\frac{dQ_1}{dQ_0}dz_1 + \ldots$$
$$+ (1+r)^{1-K} \int_{z_1} \ldots \int_{z_{K-1}} d_{K-1}(Q_{K-1}, z_{K-1})\frac{dQ_{K-1}}{dQ_0}dz_{K-1} \ldots dz$$
$$- (1+r)^{-K} \int_{z_1} \ldots \int_{z_K} d_K(Q-Q_0- \ldots -Q_{K-1},z_k)$$
$$\left(1+\frac{dQ_1}{dQ_0}+ \ldots +\frac{dQ_{K-1}}{dQ_0}\right)dz_K \ldots dz_1$$
$$=0.$$

By the induction hypothesis, however,

(xi) $$d_1(Q_1,z_1) = (1+r)^{-1} \int_{z_2} d_2(Q_2,z_2)\,dz_2 = \ldots$$
$$= (1+r)^{1-K} \int_{z_2} \ldots \int_{z_K} d_K(Q_K,z_K)\,dz_K \ldots dz_2.$$

Substituting (xi) into (x) and noticing that most of (x) cancels and that $Q_K = Q - Q_0 - \ldots - Q_{K-1}$, we get

(xii) $p_0 = (1+r)^{-K}E(p_K)$.

By combining (xii) with the induction hypothesis (vi), the proof of (vii) is completed.

HARVARD UNIVERSITY