

Intertemporal Depletion of Resource Sites by Spatially Distributed Users

By GÉRARD GAUDET, MICHEL MOREAUX, AND STEPHEN W. SALANT*

Mineral deposits and other exhaustible resources are scattered around the globe—as are their users. Yet the theory for analyzing depletion of such resources assumes that all resource sites and all users are located in the same place. When applying the theory, therefore, one is forced to disregard what often seems to be an important real-world consideration. Our purpose here is to generalize Harold Hotelling's (1931) theory of exhaustible resources to accommodate situations in which resource sites and their users are spatially distributed.

Our analysis has many applications. It can be applied to a conventional exhaustible resource like oil. Alternatively, it can be applied to an unconventional exhaustible resource like scarce landfill space.¹ Shipments of solid waste to

landfills in other states or countries have become a prominent public policy issue (Eduardo Ley et al., 2000). Finally, our analysis can be applied in situations where the “space” reflects not geographical location but characteristics of the resource (e.g., the sulfur content, hardness, and so forth, of coal) which users find distinctive.²

Only two theoretical papers in resource economics have introduced spatial considerations.³ Jean-Jacques Laffont and Michel Moreaux (1986) studied the extraction of resources from sites located along a line segment with all users at one end. Charles Kolstad (1994) was the first to consider spatially distributed users. In his formulation, consumers are distributed uniformly along a line segment with resource sites at each end. Kolstad's approach, although innovative, requires that the set of users be parti-

* Gaudet: Département de Sciences Économiques and Centre de recherche et développement économique (CRDE), Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Québec H2V 2Z9, Canada (e-mail: gerard.gaudet@umontreal.ca); Moreaux: LEERNA and IDEI, Institut Universitaire de France, Université de Toulouse I, Place Anatole France, 31042 Toulouse, France (e-mail: mmoreaux@toulouse.inra.fr); Salant: Department of Economics, University of Michigan, Ann Arbor, MI 48109 (e-mail: ssalant@umich.edu). This paper was written in part while Gérard Gaudet and Stephen Salant were both visiting researchers at INRA-Toulouse. Michel Moreaux thanks the Fondazione ENI Enrico Mattei for financial support. Gérard Gaudet also benefited from financial support from the Fonds FCAR du Gouvernement du Québec and the Social Sciences and Humanities Research Council of Canada. We are grateful to Charles Kolstad, Kai-Uwe Kühn, Jean-Jacques Laffont, Lars Olson, and Gilles Rotillon for their comments on an earlier version.

¹ When a city transports solid waste to a particular landfill, the reader should envision the landfill as instead extracting the displaced volume of space and shipping it to that city. Space in existing landfills can be augmented either by expanding old ones or creating new ones. However, because these activities are costly, the theory of exhaustible resources is still applicable, as noted by Ni-Bin Chang and Richard E. Schuler (1990), F. Dunbar and M. Berkman (1990), Mark J. Ready and Richard C. Ready (1995), and many others. All of these papers use the Hotelling (1931) model to analyze landfill depletion but none takes account of the spatial dimension of the trash trade.

² Our generalization may have other important applications as well. Recently, the standard Hotelling model was used to determine how intensively different antibiotics should be used when the usage of each “depletes” its effectiveness in killing a single type of bacteria [see Gardner Brown and David F. Layton (1996) and Brown and Ramanan Laxminarayan (2001)]. To be more relevant, this important work needs to be generalized to the case of “multiple types of bacteria.” But this involves precisely the generalization developed here.

³ There have been several complex computerized models that do take account of spatial dimensions, although they are too large to shed much light on the general principles underlying the socially efficient allocation. For example, William D. Nordhaus (1973) introduced spatial considerations into his study of the socially efficient exploitation of the world's energy resources. More recently, Ujjayant Chakravorty et al. (1997) simulated an energy model in which the distance between the final users and the resources refers to product space rather than actual distance. Resources and users are assumed to share one location but various resources (oil, coal, natural gas, or solar energy) can satisfy different end uses (electricity, heating, transportation) after a conversion cost is paid, which depends on the particular resource and end use. These costs of conversion play the same role as the cost of transporting the resource to the user in the spatial interpretation: for example, one must go to “greater lengths” to convert coal to transportation than to convert oil to transportation but can convert either resource to heating with equal ease.

tioned between the two sites at every instant. This formulation does not generalize to other spatial configurations and hence is of little use in applications.

Unlike Kolstad, we focus not on resource sites but on the final users (we call them "cities"). This change in perspective simplifies the problem and permits us to analyze the general case where cities and the resource sites they utilize can be located *anywhere*. We extend our analysis to take into account that extraction may involve "setup costs" before it can begin.⁴

We show that three basic principles of social efficiency must hold whether or not there are setup costs: each city (1) never uses more than one site at any given time; (2) always uses that site, among those that are set up, with the lowest full marginal cost (defined as including not only the per-unit cost of extraction and transport but also the imputed per unit cost of having one less unit of the resource); and (3) never reduces its usage when it switches sites.

In addition to these basic principles, three characteristics of the optimal allocation are particularly striking. First, in the absence of setup costs, the per-unit cost of the successive sites used by any given city must form an increasing sequence—the multicity generalization of the one-city result of Orris C. Herfindahl (1967). Second, if one site can serve each of two cities more cheaply than another site but one city can switch to the other site at a smaller increase in cost, then efficiency requires that this city be the first to switch to the more costly site—an application of the principle of comparative advantage.⁵ As a timely illustration, recall that New York City plans to ship its solid waste to Virginia now that its local facility, Fresh Kills, is

closing. This plan seems on its face inefficient because there are landfills in New Jersey with spare capacity. Our analysis shows when such a plan can be efficient. Finally, in the presence of setup costs, social efficiency may require that a site that is partially drawn down be abandoned by *every* city temporarily before some city returns to it. This last result requires at least three resource sites and hence cannot arise in Kolstad's two-site configuration with a mine at each end of a line segment.

In the next section, we describe the general problem when many users and many sites of the resource are arbitrarily distributed in the plane. In Section II, we apply these conditions to describe the optimal intertemporal and spatial allocation when there are no setup costs. In Section III, we extend the discussion to the case with setup costs. Section IV suggests several extensions and concludes the paper.

I. The General Problem

There are n resource sites and m cities located anywhere in the plane. The cities can be considered as m different potential users of the finite volume of resources available at each of the n sites. The stock of resources available at these sites can be thought of as stocks of fossil fuels, minerals, or landfill space.

Assume that the marginal cost of extracting resources from site i and transporting it to city k is constant and denote it c_{ik} . Further assume, to eliminate even the possibility of indeterminacies over time, that for any city no two sites have the same marginal cost. The setup cost of site i will be denoted $F_i \geq 0$. Given that we will focus on the optimization problem of a social planner, all these costs are to be interpreted as social costs, and hence are inclusive of any external costs generated by the setting up, extraction, and transportation of the resource from the site to the city.

We will denote by $S_i(t)$ the remaining resource stock at site i at time t and by $S_i^0 > 0$ the given initial reserves at that site.⁶ The quantity

⁴ In the nonspatial context, account of setup costs has been taken by Murray C. Kemp and Ngo Van Long (1984) and John M. Hartwick et al. (1986). Martin L. Weitzman (1976) also proposed a method for solving the problem of minimizing the present discounted cost of supplying a fixed flow of a depletable resource from many deposits, when the average cost of extraction from a deposit depends on how much has already been taken out of it. A positive setup cost with constant marginal cost of extraction can be viewed as a special case of such a cost configuration.

⁵ Ujjayant Chakravorty and Darrell L. Krulce (1994) deserve credit for being the first to demonstrate an analogous result in a nonspatial example with two resources and two end uses under one assumed ordering of costs.

⁶ The initial reserves at a site may be infinite, in which case it should be interpreted as a spatially located "backstop," for example, solar energy as a backstop alternative to fossil fuel exploitation or incineration as a backstop alternative to landfilling.

of the resource allocated to city k from site i at date t will be denoted $q_{ik}(t) \geq 0$. The total quantity of the resource consumed by city k at date t is therefore $q_{\bullet k}(t) = \sum_{i=1}^n q_{ik}(t)$ and the total quantity consumed by all cities from site i is $q_{i\bullet}(t) = \sum_{k=1}^m q_{ik}(t)$. The total utility derived by city k from its resource consumption at date t is $U_k(q_{\bullet k}(t))$. Total utility is assumed to be a strictly increasing, strictly concave function with $U_k(0) = 0$ and $U'_k(0)$ finite.⁷ The social rate of discount will be denoted r and τ_i will denote the date at which the setup cost is incurred for site i .

To allocate the resource stocks over cities and over time in an efficient manner, the planner will:

$$\max_{\{q_{ik}(t)\}, \tau_i, T} \int_0^T e^{-rt} \sum_{k=1}^m \left[U_k(q_{\bullet k}(t)) - \sum_{i=1}^n c_{ik} q_{ik}(t) \right] dt - \sum_{i=1}^n e^{-r\tau_i} F_i,$$

subject to

$$(1) \quad \dot{S}_i(t) = - \sum_{k=1}^m q_{ik}, \quad i = 1, \dots, n$$

⁷ In the case of landfills, the utility function represents costs saved by not having to rely entirely on recycling (or source reduction). Assume the waste stream of city k (denoted w_k) cannot be reduced. Instead it must be either (1) recycled or (2) landfilled somewhere. Denote the total cost of recycling at rate z as $C_R(z)$, where this function is strictly increasing and strictly convex with $C_R(0) = 0$ and $C'_R(w_k)$ finite and strictly positive. As a simplification to eliminate corner solutions, assume $C'_R(0) = 0$. Then we can interpret $U_k(q_{\bullet k}(t))$ as the recycling costs saved by landfilling at rate $q_{\bullet k}(t) \leq w_k$, so that recycling occurs at the reduced rate $w_k - q_{\bullet k}(t)$: $U_k(q_{\bullet k}(t)) = C_R(w_k) - C_R(w_k - q_{\bullet k}(t))$. This utility function is defined for $q_{\bullet k}(t) \in [0, w_k]$ but given our assumption that $C'_R(0) = 0$, the planner will always confine himself to $q_{\bullet k}(t) \in [0, w_k]$. Note that for this interval the utility function satisfies the properties assumed throughout: it is strictly increasing, strictly concave, passes through the origin and has a finite, strictly positive first derivative at zero. In the optimal allocation, the planner recycles city k 's waste to the point where the marginal cost of recycling additional waste equals the full marginal cost of landfilling it at the least costly of the n sites. For further discussion and the generalization that incorporates source reduction, see Gaudet et al. (1999).

$$(2) \quad q_{ik}(t) \geq 0, \quad q_{ik}(t) = 0 \quad \forall t < \tau_i,$$

$$S_i(0) = S_i^0, \quad S_i(T) \geq 0, \\ i = 1, \dots, n \quad k = 1, \dots, m.$$

Call this problem P . It can be decomposed into two subproblems. Subproblem $P1$ consists in solving problem P for given $\tau_i \geq 0$'s. Subproblem $P2$ then consists in solving

$$\max_{\tau_i \geq 0} J(\tau_1, \tau_2, \dots, \tau_n)$$

where $J(\tau_1, \tau_2, \dots, \tau_n)$ is the optimal value of the program generated by $P1$. Given the stationarity of the exogenous functions, it is optimal to incur at least one setup cost at date 0 as long as it is optimal ever to use some resource site. Hence $\tau_i = 0$ for one or more values of $i, i = 1, \dots, n$.

The current value Hamiltonian associated with Problem $P1$ is:

$$(3) \quad H = \sum_{k=1}^m \left[U_k(q_{\bullet k}(t)) - \sum_{i=1}^n [c_{ik} + \lambda_i(t)] q_{ik}(t) \right].$$

It measures the total social value derived by all the cities from consuming the resource at date t , net of all the associated costs, including the total imputed value of the resource being depleted. The variable $\lambda_i(t)$ denotes the imputed value of a unit of resource at site i at time t .

Problem $P1$ requires that for $i = 1, \dots, n$ and $k = 1, \dots, m$ the following conditions be satisfied for all $t \geq \tau_i$:

$$(4) \quad q_{ik}(t) \geq 0,$$

$$U'_k(q_{\bullet k}(t)) - c_{ik} - \lambda_i(t) \leq 0,$$

$$[U'_k(q_{\bullet k}(t)) - c_{ik} - \lambda_i(t)] q_{ik}(t) = 0$$

$$(5) \quad \lambda_i(t) = \lambda_i e^{rt}.$$

In addition, at the terminal date T , the following transversality conditions must hold:

$$(6) \quad H(\cdot)|_T = 0$$

$$(7) \quad \lambda_i(T)S_i(T) = 0, \quad \lambda_i(T) \geq 0,$$

$$S_i(T) \geq 0, \quad i = 1, \dots, n.$$

The economic interpretation of these necessary conditions merits discussion. Condition (4) is necessary to maximize at each date the Hamiltonian associated with problem *PI*. If city k is using site i at date t , then condition (4) requires that the marginal utility derived from using site i be equal to the *full* marginal cost of doing so. This “full marginal cost” includes not only the costs of extracting the resource from site i and transporting it to city k , but also the imputed cost $[\lambda_i(t)]$ of depleting an extra unit of the resource at site i . This imputed cost takes account of the net utility forgone in not being able to consume that marginal unit of the resource at a future date—perhaps at a *different* city. Condition (4) also says that if the full marginal cost of city k using site i at date t strictly exceeds the marginal benefit, city k should not use that site $[q_{ik}(t) = 0]$ at that date.

Condition (5) is nothing but the Hotelling rule of nonrenewable resource exhaustion (Hotelling, 1931) applied to each resource site i , $i = 1, \dots, n$. It says that the shadow value assigned to the marginal unit of the resource remaining at the site must grow at the discount rate r .

The terminal conditions (6) and (7) serve to determine the optimal values of T , the date at which all the resource sites are exhausted, and of the $\lambda_i(T)$'s. Condition (6) requires that the value of the Hamiltonian (H)—which measures the total surplus net of the opportunity cost of the resource being depleted—be zero at the terminal date T . If instead it were strictly positive, then the depletion program could be improved upon by delaying somewhat the terminal date.⁸ By substitution for the $\lambda_i(T)$'s from (4), one verifies that, because marginal cost is independent of the remaining resource stock and the utility functions are strictly concave, condition (6) implies that no resource site be operating at a positive rate at the terminal date [i.e., $q_{ik}(T) =$

0 for all $i = 1, \dots, n$]. Condition (7) says that the value of the remaining reserves must be zero at the terminal date T . Again, in that marginal cost at any date t is independent of the remaining reserves, this will require that all sites be exhausted at the terminal date, with $\lambda_i(T) > 0$ being determined by the condition that $S_i(T) = 0$ be satisfied.

Three basic results follow directly from conditions (4) and (5). First, *it is never optimal for a city to use more than one site over any interval of time.*⁹ For suppose city k were using simultaneously both site i and site j over some interval $\Delta t > 0$ ending at time t . Since this means that $q_{ik}(t) > 0$ and $q_{jk}(t) > 0$ over that interval, conditions (4) and (5) imply that $[U'_k(q_{\bullet k}(t)) - c_{ik}]e^{-rt} = [U'_k(q_{\bullet k}(t - \Delta t)) - c_{ik}]e^{-r(t - \Delta t)}$ and $[U'_k(q_{\bullet k}(t)) - c_{jk}]e^{-rt} = [U'_k(q_{\bullet k}(t - \Delta t)) - c_{jk}]e^{-r(t - \Delta t)}$. Subtracting the second of these equations from the first, we find that they imply $c_{ik} - c_{jk} = (c_{ik} - c_{jk})e^{r\Delta t}$. But this is impossible because, by assumption, r is positive and $c_{ik} \neq c_{jk}$ for all $i \neq j$.

Second, *among all the resource sites that are already set up, each city will, at any given date, use only the one with lowest full marginal cost.* This result follows given that maximization of H at each t with respect to $q_{ik}(t)$ requires that, for each k , we maximize

$$H_k = U_k(q_{\bullet k}(t)) - \sum_{i=1}^n [c_{ik} + \lambda_i(t)]q_{ik}(t),$$

with respect to $q_{ik}(t)$, subject to $\sum_{i=1}^n q_{ik}(t) = q_{\bullet k}(t) \geq 0$. This maximization can be achieved by first maximizing H_k over $q_{ik}(t) \geq 0$ subject to some specified $q_{\bullet k}(t)$ and then maximizing over all $q_{\bullet k}(t) \geq 0$. Given the linearity of the full total cost and the linearity of the constraint, which implies that the sites are perfect substitutes in use, the first maximization always requires that for all t and for all k , $q_{ik}(t)$ be set to zero except for the one for which the full marginal cost $c_{ik} + \lambda_i(t)$ is smallest among all the sites that are set up.

⁹ This result is a consequence of our assumption of linear technologies. If sites were instead capacity constrained or had strictly convex cost functions, then it could be efficient for the same city to use more than one site over an interval of time.

⁸ Because we assume $U'_k(0)$ to be finite for all $k = 1, \dots, m$, then T will necessarily be finite.

Third, when a city switches from one resource site to another, its consumption of the resource does not jump down. Indeed, as we have just shown, if city k switches from site i to site j at some date τ , then it must be the case that $c_{jk} + \lambda_j(\tau^+) \leq c_{ik} + \lambda_i(\tau^-)$. Otherwise, the city would be switching to a site with a strictly larger full marginal cost, in violation of the previous result. Intuitively, the planner would assign a different site to a given city only if one became available with a (weakly) cheaper full marginal cost. It follows from condition (4) that $U'_k(q_{\cdot k}(\tau^+)) \leq U'_k(q_{\cdot k}(\tau^-))$. By the concavity of the utility function, this means that $q_{\cdot k}(\tau^+) \geq q_{\cdot k}(\tau^-)$.

The solution to problem $P2$ will determine the optimal setup dates for each resource site. We leave the discussion of the necessary conditions for problem $P2$ to Section III, where we discuss the consequences of allowing for setup costs. The case where there are no setup costs, to which we now turn, is equivalent to a situation where all sites are already set up at $t = 0$. In that case, those necessary conditions are trivially satisfied.

II. No Setup Costs

If the cost of setting up sites is negligible ($F_i = 0$), then it is optimal to set up every site at the outset and we can focus on $P1$. In the absence of setup costs, the objective function in $P1$ is strictly concave in the rates of shipment from site i to city k and any program satisfying both the equation of motion [equation (1)] and conditions (5)–(7) must be optimal.

The following algorithm can be used to solve the entire set of first-order conditions iteratively: (1) assign to each site an initial multiplier and let each multiplier subsequently grow at the exogenous rate of interest; (2) for each of the m cities, assemble the set of the n full marginal costs available at each date and assign to that city at that date the site with the smallest full marginal cost; (3) assume that each city uses the designated site to the point where the marginal benefit from additional usage equals the full marginal cost; (4) determine the cumulative usage of each site over time and across cities and compare the cumulative usage to the initial reserves of that site; (5) if, for each site, cumulative usage exactly matches the initial stock, then the optimum has been identified; (6) otherwise,

the multipliers must be revised and the process repeated. A gradient algorithm will converge to the unique solution. It is straightforward to verify that every first-order condition will then be satisfied.

In the optimal program, resources with different marginal costs may be utilized at the same time. As Kolstad (1994) points out, this contrasts with the one-city case where no site with a higher per-unit cost is utilized until every site with a lower per-unit cost is exhausted (Herfindahl, 1967).

A. The Generalized Herfindahl Principle

Clearly, no aggregate statement about the order of extraction can be made in the multicity case. However, Herfindahl's insight continues to apply at the city level. That is, each city will use resource sites in the order of their marginal costs. Hence, if city k ever uses sites i and j and $c_{jk} > c_{ik}$, then it must complete its usage of site i before it begins its usage of site j . The proof mirrors that in the one-city case. There are two logical possibilities: either $\lambda_i \leq \lambda_j$ or $\lambda_i > \lambda_j$. In the former case, the full marginal cost of using site i is always smaller for city k than the corresponding cost of using site j and city k would never use site j , contradicting the hypothesis that city k eventually uses both sites. Hence, $\lambda_i > \lambda_j$. If city k uses one site and then the other, there will come a date (t_k), defined by $\lambda_j e^{rt_k} + c_{jk} = \lambda_i e^{rt_k} + c_{ik}$, when city k is indifferent between the two sites. Subsequent to that date, city k must use site j because its full marginal cost (the left-hand side above) is smaller; prior to that date, city k must use site i because its full marginal cost (the right-hand side above) is smaller. It may not be optimal for some city to use one or more of the n sites. However, those sites utilized by any given city must be utilized in order of the marginal cost of serving that city.

B. The Principle of Comparative Advantage

The generalized Herfindahl principle requires that if sites i and j serve cities A and B eventually and site i can serve each city at a smaller cost, then each city will use the cheaper site (site i) first. The principle is, therefore, consistent with (1) both cities switching from site i to site j at the same time or, alternatively, (2) either of the

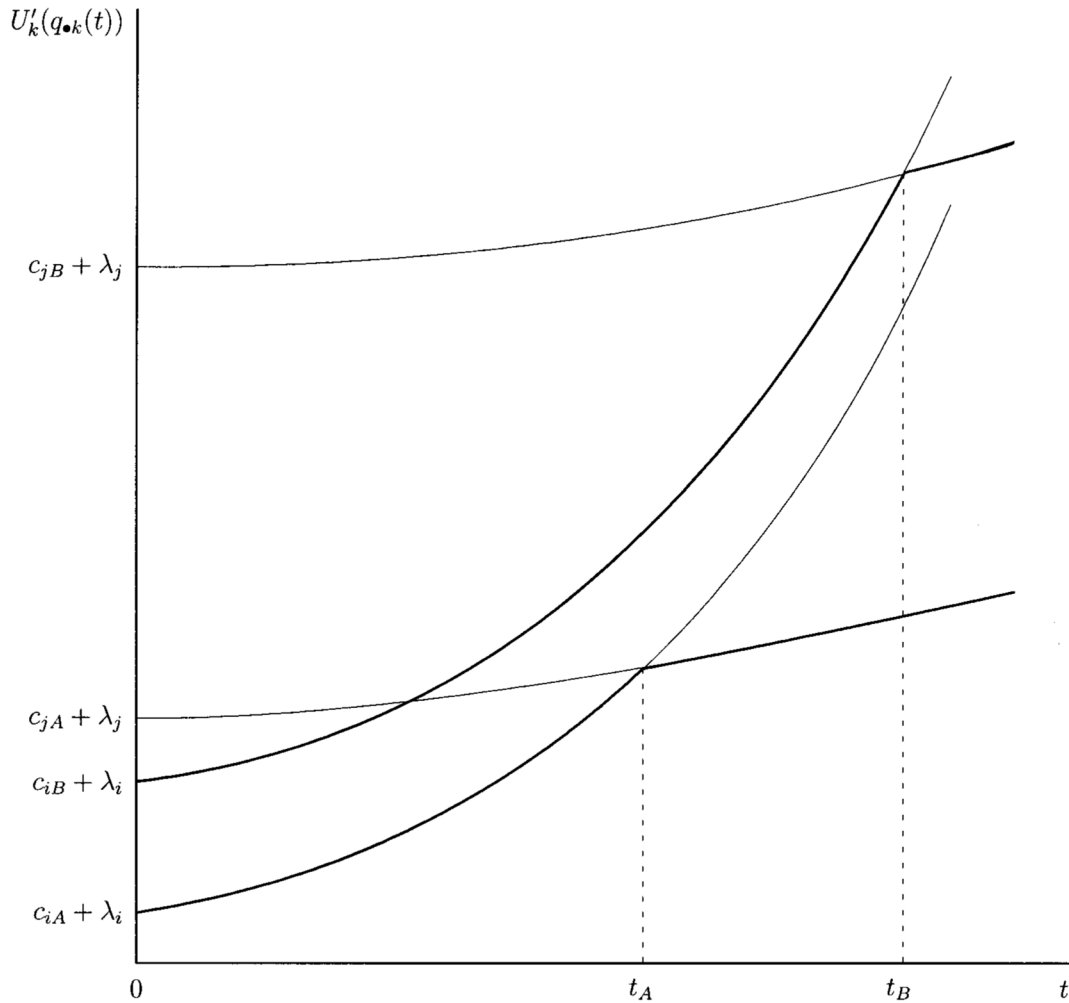


FIGURE 1. CITY A SWITCHES TO A MORE COSTLY SITE BEFORE THE CURRENT ONE IS EXHAUSTED

two cities being the first to switch from site i to site j .

The principle of comparative advantage supplements the generalized Herfindahl principle by further restricting what can be optimal. It determines which city is the first to switch to the higher-cost site. Suppose $0 < c_{jA} - c_{iA} < c_{jB} - c_{iB}$ so that social costs increase more if city B (rather than city A) uses a unit from site j instead of site i . Then site j has a *comparative advantage* in serving city A , even if it has no *absolute advantage*. In such cases, it is efficient for city A to switch to site j , even though city B remains with site i . To verify this, let t_A denote

the date when city A switches from site i to site j . At that date, $(\lambda_i - \lambda_j)e^{rt_A} = c_{jA} - c_{iA}$. But because $c_{jA} - c_{iA} < c_{jB} - c_{iB}$, city B will still utilize site i because it has a smaller full marginal cost than site j , even though city A is switching to site j . The situation is depicted in Figure 1.

That it is optimal for city A to utilize the more costly site, even though space remains in the less costly one may still seem counterintuitive. We conclude, therefore, by verifying directly that this assignment cannot be improved upon.

At t_A ,

$$(8) \quad c_{iA} + \lambda_i e^{rt_A} = c_{jA} + \lambda_j e^{rt_A},$$

and city A begins using site j instead of site i . At $t_B > t_A$,

$$(9) \quad c_{iB} + \lambda_i e^{rt_B} = c_{jB} + \lambda_j e^{rt_B},$$

site i is completely depleted, and city B begins using site j . Therefore, for $t \in [t_A, t_B]$, city A is using site j while resources are still available at site i and $c_{iA} < c_{jA}$.

Transferring a unit of consumption from site j to site i at $t \in [t_A, t_B]$ results in a cost saving of $e^{-rt}(c_{jA} - c_{iA}) > 0$. However, because site i will be exhausted by city B at t_B , this transfer must be offset by a similar transfer from site i to site j by city B at some date $(t + \Delta t) \in [t, t_B]$. This last transfer will result in a cost increase of $e^{-r(t+\Delta t)}(c_{jB} - c_{iB})$. Hence, the resulting overall cost saving, in present value terms, is $e^{-rt}[(c_{jA} - c_{iA}) - e^{-r\Delta t}(c_{jB} - c_{iB})]$, which is nonpositive.¹⁰ Therefore the program cannot be improved upon by such a transfer and it is indeed optimal for city A to switch to the higher-cost resource site j before the lower cost site i is fully used up.

III. Extension to the Case of Setup Costs

Readying resource sites before they can be extracted is often a costly undertaking. To broaden the applicability of our analysis, we now take account of setup costs. The analysis is more complex because we must derive and examine the conditions that hold when problem $P2$ is solved.

Let $M_i(\tau_i) \subseteq \{1, 2, \dots, m\}$ denote the subset of cities that switch to site i when it is set up at τ_i and let $h(k) \in \{1, 2, \dots, n\}$, $h(k) \neq i$, denote the resource site used by city k just before switching to site i . Then the following conditions must be satisfied at τ_i :

$$(10) \quad \sum_{M_i} G_k(q_{h(k)k}(\tau_i^-)) + rF_i = \sum_{M_i} G_k(q_{ik}(\tau_i^+))$$

¹⁰ Replacing the cost differences in the text by differences in shadow values (appropriately discounted) using conditions (8) and (9), we can rewrite the cost saving as $e^{-rt}(e^{rt_A} - e^{r(t_B - \Delta t)})(\lambda_i - \lambda_j) \leq 0$.

and

$$(11) \quad \sum_{M_i} G'_k(q_{h(k)k}(\tau_i^-)) \dot{q}_{h(k)k}(\tau_i^-) \leq \sum_{M_i} G'_k(q_{ik}(\tau_i^+)) \dot{q}_{ik}(\tau_i^+),$$

where $G_k(q_{ik}) = U_k(q_{ik}) - U'_k(q_{ik})q_{ik}$. Because when site i is used by city k , $U'_k(q_{ik}) = c_{ik} + \lambda_i e^{rt}$, we can rewrite this as $G_k(q_{ik}) = U_k(q_{ik}) - (c_{ik} + \lambda_i e^{rt})q_{ik}$. Hence, it is the gross utility that city k obtains from consuming q_{ik} units of the resource from site i net of the full cost (including the imputed cost) of extracting and transporting it from site i .¹¹

Condition (10) is a first-order condition. It implies that it is optimal to delay setting up site i if the interest saved by postponing the setup cost is strictly greater than the loss in net consumer surplus from having to use the old sites a little longer. In making this calculation, the planner must take into account the loss in consumer surplus at every city that would have switched to site i if it were opened at τ_i .

Condition (11) is a second-order condition. Because $G'_k(q_{ik}) = -U''_k(q_{ik})q_{ik}$ and, by the necessary conditions to problem $P1$, $U''_k(q_{ik})\dot{q}_{ik} = \dot{\lambda}_i = r\lambda_i$, this condition is best rewritten as

$$(12) \quad \sum_{M_i} \lambda_{h(k)}(\tau_i^-) q_{h(k)k}(\tau_i^-) \geq \lambda_i(\tau_i^+) \sum_{M_i} q_{ik}(\tau_i^+).$$

It says that the imputed value of the total quantity of the resource being consumed just after the switch to site i must be no greater than the

¹¹ Condition (11) must hold in order for the Hessian of $J(\tau_1, \tau_2, \dots, \tau_n)$ to be locally negative semidefinite. Condition (10) must hold in order for the derivatives of $J(\tau_1, \tau_2, \dots, \tau_n)$ with respect to τ_i , $i = 1, \dots, n$, to be zero. Notice that the first-order condition for the corner case at $\tau_i = 0$ can be obtained by replacing the first term on the left-hand side of condition (10) by zero and its equality sign by a weak inequality sign, indicating that the right-hand side is at least as large as the left-hand side. That is, if it is optimal to open site i at $t = 0$, then the net benefit from opening it must be at least as large as the interest that would be saved by delaying its opening.

total imputed value of that consumed just before the switch.

With only one city and many resource sites, the conclusions of Hartwick et al. (1986) apply immediately. As in the case without setup costs, the city must exploit the resource sites in strict sequence: each site must be exhausted before going on to the next. However, unlike the case without setup costs, sites need not be exploited in order of their marginal costs; it may be optimal, for example, to postpone usage of a site with the lowest marginal cost if opening it involves relatively high setup costs.

It is easy to see that opening sites in order of their marginal costs is no longer necessary for an optimum in the presence of setup costs. For suppose the lone city is city k and that it switches at τ_j from site i to site j , which requires that a setup cost of $F_j > 0$ be incurred at that date. Then, remembering that it never uses more than one site, we deduce from the first-order condition (10), which must hold at τ_j , that $q_{jk}(\tau_j^+) > q_{ik}(\tau_j^-)$: the city's resource consumption jumps up. The second-order condition (11), which must also hold at τ_j , can be written

$$\begin{aligned} & (\lambda_i(\tau_j^-) - \lambda_j(\tau_j^+))q_{ik}(\tau_j^-) \\ & \geq \lambda_i(\tau_j^+)(q_{jk}(\tau_j^+) - q_{ik}(\tau_j^-)). \end{aligned}$$

Because the right-hand side is strictly positive, so must be the left-hand side, and hence $\lambda_i(\tau_j^-) > \lambda_j(\tau_j^+)$. Now the city's consumption path must also satisfy, at all t , the first-order condition (4), and, as a result, $c_{jk} + \lambda_j(\tau_j^+) < c_{ik} + \lambda_i(\tau_j^-)$. This does not require $c_{jk} > c_{ik}$.¹²

In the case of many cities and many resource sites with setup costs, the three basic results established in Section I still hold: each city never uses more than one site at any given time, always uses the one with the lowest full marginal cost among all those that are set up, and never reduces its usage when it switches sites. But they admit a very remarkable possibility as well. Not only can it be optimal for a city to

exploit a high marginal cost site before a low marginal cost one (because of the presence of fixed costs) or to abandon a low marginal cost site in favor of a higher cost one while the low cost one is still available (because there are other cities using the sites whose behavior has an impact on their full marginal cost) but a city may well abandon a site and come back to it at a later date. In optimizing, the planner may thus appear to vacillate.

Clearly, this requires that the city switch from a higher-cost site to a lower-cost site when it abandons the site originally and then from a lower-cost site to a higher-cost site when it resumes usage of the abandoned site. This can never happen in the case of a single city with or without setup costs because, in all such cases, a site must be fully depleted before it is abandoned and there can be no point in returning to it. Nor can this ever occur with multiple cities in the absence of setup costs. For, although it is possible for a city to switch to a higher-cost site while a lower-cost site is still available, it is never optimal to return to the low-cost site because transitions for any given city must always be to sites with higher marginal costs. The possibility requires both multiple cities and setup costs. As we have discussed, when there are multiple cities it is sometimes optimal for a city to abandon a site, even though that site is not fully depleted. When there are also setup costs, the site that city abandons may be a high-cost site, which has been exploited before a low-cost one has been set up. Once the lower cost site is set up, however, the city must switch to it because it would be, among the sites set up, the one with the lowest full marginal cost. Once that site is exhausted, however, a return by that city to the abandoned site would be a standard transition to a site of higher marginal cost.

This counterintuitive result is best understood with the help of an example. Assume there are two cities (cities A and B) and three resource sites (sites 1, 2, and 3). For the sake of the discussion, we may assume that the cost of transporting a unit of the resource from a particular site is strictly proportional to the distance from the site to the city and that, other than the setup cost, this is the only cost. Assume that sites 1 and 2 have no setup cost ($F_2 = F_1 = 0$) but that site 3 has a very large setup cost ($F_3 > 0$). Site 2 is small, whereas sites 1 and 3 are

¹² Recall that in the one-city case with no setup cost, the first-order condition (4) implies $c_{jk} + \lambda_j(\tau_j) = c_{ik} + \lambda_i(\tau_j)$, which requires $c_{jk} > c_{ik}$ because $\lambda_i(\tau_j) > \lambda_j(\tau_j)$ by condition (12). In this case, τ_j is the date at which the city switches from site i to site j at zero setup cost because all the sites are assumed already set up at the outset.

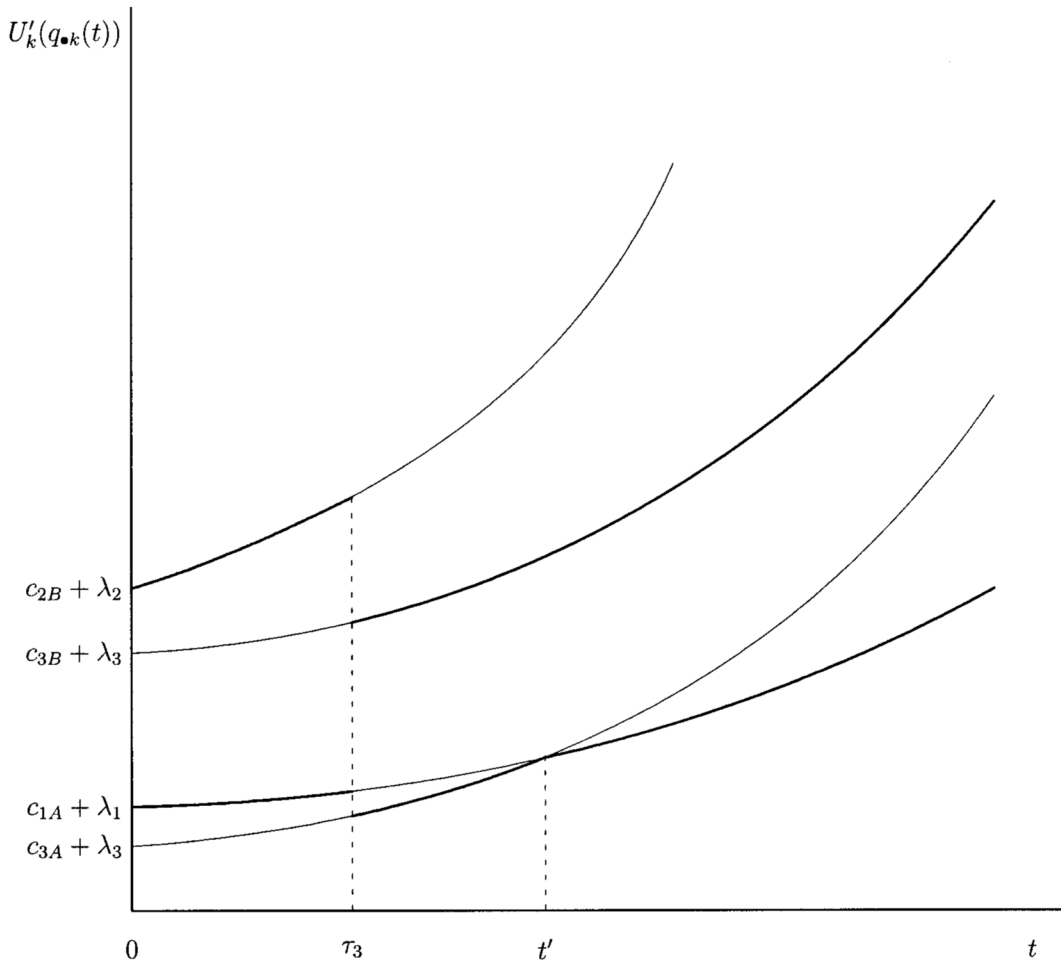


FIGURE 2. WITH SETUP COSTS, SITE 1 IS ABANDONED AND LEFT COMPLETELY IDLE UNTIL CITY A RETURNS TO EXHAUST IT

larger. Site 1 is closer to city *A* than it is to city *B* ($c_{1A} < c_{1B}$) and it is closer to city *A* than is site 2 ($c_{1A} < c_{2A}$). Site 3 is located at city *A* ($c_{3A} = 0$), whereas city *B* is equidistant from sites 2 and 3 ($c_{2B} = c_{3B}$).

Now pick the shadow values of the resource stocks such that $\lambda_2 > \lambda_3 > \lambda_1$, $c_{1A} < \lambda_3 - \lambda_1$ and $c_{1B} - c_{2B} > \lambda_1$. We then have $c_{2A} + \lambda_2 > c_{1A} + \lambda_1 > c_{3A} + \lambda_3$ and $c_{1B} + \lambda_1 > c_{2B} + \lambda_2 > c_{3B} + \lambda_3$. The resulting time paths of marginal utility at each city are depicted in Figure 2.

Given the fact that sites 1 and 2 are already set up and that site 3 has a very large setup cost, at first city *A* uses site 1 and city *B* uses site 2. Site 2 is relatively small and is exhausted by

city *B* at some date τ_3 , whereas site 1 still has some reserves left. However, because the cost of transporting the resource from site 1 to city *B* is very high, it becomes optimal to set up site 3, even though this means incurring a setup cost of F_3 . Once set up, it becomes advantageous to have city *A* switch to site 3 as well, given that the full marginal cost of site 3 to city *A* is smaller than that of site 1 ($c_{1A} + \lambda_1(\tau_3) > c_{3A} + \lambda_3(\tau_3)$). Because $\lambda_3(t)$ and $\lambda_1(t)$ both grow at the same constant rate and $\lambda_3 > \lambda_1$, $c_{3A} + \lambda_3(t)$ will eventually cross $c_{1A} + \lambda_1(t)$ from below at, say, date t' . It therefore becomes optimal for city *A* to return to site 1 at t' because beyond that date $c_{1A} + \lambda_1(t) < c_{3A} + \lambda_3(t)$.

Along each path just described, resource consumption is given by equalizing marginal utility at each city to the full marginal cost of the site being used, calculated using the shadow values picked initially. The necessary condition (4) is therefore satisfied. So is the necessary condition (5), by having the shadow values grow at the same constant rate of discount. The switching date τ_3 is chosen to satisfy the necessary condition (10). As for the second-order condition (12), which requires that the imputed value of the total resource stock being consumed not jump up when a switch occurs, it may be written

$$\begin{aligned} & (\lambda_3(\tau_3^+) - \lambda_1(\tau_3^-))q_{3A}(\tau_3^+) \\ & + (\lambda_3(\tau_3^+) - \lambda_2(\tau_3^-))q_{3B}(\tau_3^+) \\ & \leq \lambda_1(\tau_3^-)(q_{1A}(\tau_3^-) - q_{3A}(\tau_3^+)) \\ & + \lambda_2(\tau_3^-)(q_{2B}(\tau_3^-) - q_{3B}(\tau_3^+)). \end{aligned}$$

Notice that, although it constrains the λ_i 's, the fact that $\lambda_3(\tau_3^+) - \lambda_1(\tau_3^-) > 0$ and $\lambda_3(\tau_3^+) - \lambda_2(\tau_3^-) < 0$ does not prevent the condition from being satisfied. Hence, it is possible to set the multipliers and F_3 so that the necessary conditions are all satisfied and then to choose initial stocks so that the usages required in this example are exactly available.

IV. Conclusion

We have generalized the standard model of nonrenewable resource depletion to the common situation where reserves and users are located arbitrarily over space and transport costs are nonnegligible. As we have shown, a variety of new phenomena can arise. Although every city will use resources in the order of its marginal costs, it may be socially optimal for some cities to abandon a nearby resource site in preference for a more distant and more costly site, even though other cities continue to exploit the nearby site. Even though the distant site would have no absolute advantage in serving either city, it must always have a comparative advantage in serving one of them. This puts New York City's plan to ship its waste to Virginia, even though a closer site has capacity in New Jersey, in a new light: the plan is not, on its face, inefficient. We also show that, in the pres-

ence of setup costs, it may be socially optimal for every user patronizing a particular site to abandon it, even though reserves remain—and then for some users to return to it after a delay. This “vacillation result” has no counterpart either in models that ignore spatial considerations entirely or, alternatively, in those that restrict the location of mines to the ends of a line segment.

Our analysis has many applications—to conventional resources like oil, to unconventional resources like scarce landfill space (see Ley et al., 2000), and to resources whose characteristics can be envisioned as “close to” or “far from” what heterogeneous consumers regard as ideal. Our generalization will permit many applications to be addressed without having to ignore what is often an important feature of the problem.

REFERENCES

- Brown, Gardner and Laxminarayan, Ramanan.** “Economics of Antibiotic Resistance: A Theory of Optimal Use.” *Journal of Environmental Economics and Management*, 2001 (forthcoming).
- Brown, Gardner and Layton, David F.** “Resistance Economics: Social Cost and the Evolution of Antibiotic Resistance.” *Environment and Development Economics*, July 1996, 1(3), pp. 349–55.
- Chakravorty, Ujjayant and Krulce, Darrell L.** “Heterogeneous Demand and Order of Resource Extraction.” *Econometrica*, November 1994, 62(6), pp. 1445–52.
- Chakravorty, Ujjayant; Roumasset, James and Tse, Kinping.** “Endogenous Substitution among Energy Resources and Global Warming.” *Journal of Political Economy*, December 1997, 105(6), pp. 1201–34.
- Chang, Ni-Bin and Schuler, Richard E.** “Optimal Pricing of Sanitary Landfill Use Over Time.” Mimeo, Cornell University, 1990.
- Dunbar, F. and Berkman, M.** “Sanitary Landfills Are Too Cheap.” *Waste Age*, May 1987, pp. 91–9.
- Gaudet, Gérard; Moreaux, Michel and Salant, Stephen W.** “Intertemporal Depletion of Resource Sites by Spatially-Distributed Users.” Mimeo, University of Michigan, 1999.
- Hartwick, John M.; Kemp, Murray C. and Long, Ngo Van.** “Set-up Costs and the Theory of

- Exhaustible Resources." *Journal of Environmental Economics and Management*, September 1986, 13(3), pp. 212–24.
- Herfindahl, Orris C.** "Depletion and Economic Theory," in Mason Gaffney, ed., *Extractive resources and taxation*. Madison, WI: University of Wisconsin Press, 1967, pp. 68–90.
- Hotelling, Harold.** "The Economics of Exhaustible Resources." *Journal of Political Economy*, April 1931, 39(2), pp. 137–75.
- Kemp, Murray C. and Long, Ngo Van.** "Towards a More General Theory of the Order of Exploitation of Non-Renewable Resource-Deposits," in Murray C. Kemp and Ngo Van Long, eds., *Essays in the economics of exhaustible resources*. Amsterdam: Elsevier Science, 1984.
- Kolstad, Charles.** "Hotelling Rents in Hotelling Space: Product Differentiation in Exhaustible Resource Markets." *Journal of Environmental Economics and Management*, March 1994, 26(2), pp. 163–80.
- Laffont, Jean-Jacques and Moreaux, Michel.** "Bordeaux Contre Gravier: Une Analyse par les Anticipations Rationnelles," in Gérard Gaudet and Pierre Lasserre, eds., *Ressources naturelles et théorie économique*. Québec: Presses de l'Université Laval, 1986, pp. 231–53.
- Ley, Eduardo; Macauley, Molly K. and Salant, Stephen W.** "Restricting the Trash Trade." *American Economic Review*, May 2000 (*Papers and Proceedings*), 90(2), pp. 243–47.
- Nordhaus, William D.** "The Allocation of Energy Resources." *Brookings Papers on Economic Activity*, 1973, (3), pp. 529–670.
- Ready, Mark J. and Ready, Richard C.** "Optimal Pricing of Depletable, Replaceable Resources: The Case of Landfill Tipping Fees." *Journal of Environmental Economics and Management*, May 1995, 28(3), pp. 307–23.
- Weitzman, Martin L.** "The Optimal Development of Resource Pools." *Journal of Economic Theory*, June 1976, 12(3), pp. 351–64.