ON PRICING WITH LUMPY INVESTMENT:

Two Problems of Boiteux*

Glenn A. Woroch

Telecommunications Research Laboratory
GTE LABORATORIES INCORPORATED
Waltham, Massachusetts 02254

November 1985
Revised November 1987

Abstract

"Lumpy" increments to capital stock are warranted by increasing returns to plant construction. Following Boiteux, we attempt to reconcile the resulting double mismatch between capacity and demand, and between revenue and cost. Ignoring the budget problem at first, we derive conditions for optimal prices and for the optimal size and frequency of investment projects. Price is set at marginal operating cost and results in steady inflation which ends as price plummets at the start of the next investment cycle. The optimal policy is shown to create a persistent revenue shortfall which leads us to consider the budget–constrained welfare problem. Price then obeys the Ramsey–Boiteux inverse–elasticity rule at each instant. The second–best policy also implies that, roughly speaking, production is more capital intensive. Finally we evaluate the performance of policies of delegating pricing or investment control to the firm.

---

* I thank Paul Romer for his written comments on the first draft of this paper without implicating him in my errors. A recent version was presented at the Econometric Society European Meetings in Copenhagen, August 1987, and at the 14th annual conference of the European Association for Research in Industrial Economics in Madrid, September 1987.

1. INTRODUCTION

Marcel Boiteux (and his compatriots) grappled with several of the obstacles to allocative efficiency caused by technological nonconvexities. He proposed a pricing rule that minimizes welfare losses when scale economies in production preclude a balanced budget. Boiteux recognized that increasing returns also commonly occur in plant *construction*, and that they can be realized through a program of large, periodic investments. The lumpy nature of capital accumulation that results, however, prevents the ideal coordination of capacity and demand. This paper attempts to reconcile the double mismatch between capacity and demand, and between revenue and cost that stem from construction economies.

We begin by introducing scale economies in construction into a standard model of accumulation. Capital investment is taken to be perfectly divisible and fungible, and to depreciate at a constant rate over an infinite horizon. Lumpy additions to the stock of capital realize a savings over a smooth accumulation path: the firm trades off lower unit costs of construction against the higher "shadow" costs of holding too much or too little capacity. Assuming demand does not grow over time nor vary in a systematic fashion only stationary solutions to the firm's problem need be considered.

At the social optimum, price is set at marginal operating cost causing it to steadily inflate over time only to plummet as each new investment cycle begins. The optimal size and frequency of investment projects renders the firm financially insolvent. After tracing the budget problem back to construction economies, we solve the constrained welfare problem. Prices obey the familiar Ramsey–Boiteux inverse–elasticity rule at each date. The solution also reveals that, roughly speaking, the budget constraint reduces the capital intensity of the enterprise. It can even happen that a monopoly invests *more* on average than is socially optimal. An algebraic example is constructed to illustrate these points.

Turning to the question of implementation, we entertain several policies alternatives aimed at supporting the second–best solution while stopping short of complete control by a central authority. First we consider a decentralized policy of price control in which investment decisions are entrusted to the firm. It is shown that, under certain conditions, peak capacity is too large, but projects — while also too large — occur less frequently than is optimal. A second approach which instead delegates pricing to the firm does not fare much better. While neither method induces a profit–maximizing firm to behave optimally, price control appears to be more attractive if only because it better approximates the efficient investment policy.

Before launching into a description of the model, it will help to set the stage by reviewing some of the previous research in this area. Chenery (1952) first posed the problem of optimal expansion to meet growing demand when capacity was purchased in pre–determined lumps. Price played no allocative role in his original formulation, nor in the many refinements that followed. Earlier Boiteux (1949) had recognized that construction economies naturally generated indivisibilities of this sort but instead chose to concentrate on the problem of rationing demand as it varies over the load cycle. His work spawned the huge literature on peak–load pricing. In one development Williamson (1966) derived a condition for the optimal timing (but not size) of projects when demand grows over time.

In a now famous piece, Boiteux (1956a) provided a uniform–pricing solution to the budget–balancing problem in a multiproduct, static framework. Recently, Brock and Dechert (1985) imbedded the problem in a dynamic context having static scale economies in production, but not in plant construction. Mohring (1970) added a break–even constraint to the peak–load model and produced conditions for the optimal level of investment and prices. Along these lines Starrett (1978) re–considered some tentative pricing principles offered by Boiteux (1956b) to cope with lumpy investment. In a rather general set–up Starrett derived the optimal size and frequency of projects when demand grows and capital depreciates. He also developed rules for efficient pricing over a single investment cycle, and in the process, identified a condition sufficient for the enterprise to break even. We continue where Starrett left off, examining price–investment policies when this condition fails to hold. Throughout it will be clear that Starrett's treatment serves as a valuable touchstone that provides a frequent check on our results.

2. THE PRICE–INVESTMENT PROBLEM

Demand for the single, nonstorable good is represented by $P(q)$. It is stable and separable across time as if the income and population of customers were unchanging. The absence of income effects justifies the use of consumer surplus as a measure of welfare $S(q) = \int_0^q P(v)dv$. The interest rate r serves as both the private and the social rate of discount.

Production requires the services of durable plant and variable factors. Variable cost — written $V(q,k)$ — is the minimum cost of producing output q with capital stock k. We depart from the tradition in this literature (including Starrett's general treatment) of assuming unit operating costs are constant up to a fixed capacity. By taking the underlying technology to be weakly convex, V becomes convex in (q,k) since it is the value of a convex program. We have deliberately separated out the role of scale economies in *construction* from decreasing returns in *production*. Standard assumptions ensure that variable costs rise with output and fall with capacity: $V_q > 0$, $V_k < 0$. We also assume that $V_{qk} < 0$ which implies that capital usage increases with the scale, a realistic property that will be crucial to some of our results.

The firm purchases a plant of size I on the spot market at a cost of $C(I)$. It is a "turn–key" plant so that the capacity is instantly available. Importantly, capital is *infinitely divisible* so that any lumpiness arises endogenously and not as an artificial constraint.[1] As usual, plants of different vintages are *fungible* and depreciate at the constant rate δ. Repair and maintenance of capital equipment would serve to retard depreciation but this possibility is ignored. Finally, all investments are assumed to be sunk, commanding no value in alternative uses once they have been put in place.

At least at small scales, plant construction displays increasing returns. This is captured by a U–shaped average cost curve: $C(I)/I \gtrless C'(I)$ as $I \lessgtr I^0$ where I^0 is the positive (and possibly infinite) efficient scale.

---

[1]While we do not allow for smooth accumulation of capital, it is possible to show that none will occur under reasonable conditions. Construction work is progress does accumulate gradually, but this is merely *potential* capacity which becomes *usable* capacity only in lumps.

This property deserves additional motivation. Behind the scenes $C(I)$ solves a cost–minimization problem given a construction technology and given factor prices including the interest rate. Despite the close connection, the construction cost function does not automatically inherit the scale properties of the underlying technology. Increasing returns in plant production (e.g., intertemporal economies of learning by doing) certainly will tend to drive down unit construction costs. Nevertheless, decreasing returns do not imply that the unit cost of a finished plant will decline. Suppose, for example, that the interest rate is zero. The cost–minimizing program then dictates accumulation at a rate equal to the minimum efficient scale until the desired plant size is reached. In that case construction cost will exhibit constant returns to scale since unit cost is constant. At the other extreme, unit costs eventually rise when cumulative interest charges swell with long project lifetimes.

Given the regularity of cost and demand over time, nonstationary optima will be rare and so only stationary solutions to the price–investment problem are considered. If *peak capacity* at the start of a cycle immediately following investment is K, then *base capacity* of $Ke^{-\delta T}$ remains after a cycle of length T. *Project size* $I = K(1-e^{-\delta T})$ is the amount necessary to restore capacity to its peak level.[2] The sawtoothed pattern of capital stock that results is depicted in Figure 1. Early in each cycle there is an excess of capital (measured against the ideal stock for the produced level of output) followed by a period of relative scarcity of capacity.

[2]Alternatively, if capital was *not* fungible then the old plant would have to be entirely replaced with a new plant of peak size. To do so, we would replace $C(I)$ with $C(K)$ in (1) below, and so on. This approach should not alter our conclusions appreciably.

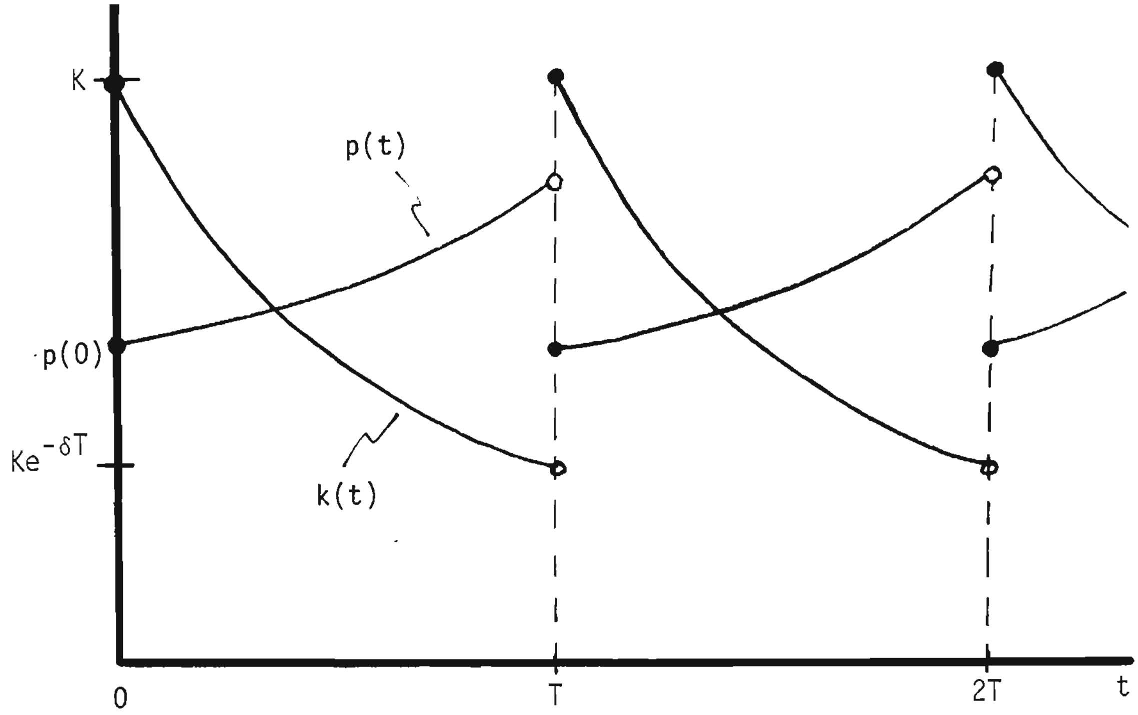

Figure 1: Stationary Cycles of Price and Capital Stock

Discounted welfare is given by:

$$(1)\quad W \equiv \frac{1}{1-e^{-rT}}\left\{\int_0^T\left[S(q(t)) - V(q(t),Ke^{-\delta t})\right]e^{-rt}dt - C(I)\right\}$$

where the factor outside the curly brackets is derived by discounting each cycle back to the starting time:

$$(2)\quad \frac{1}{1-e^{-rT}} = \sum_{n=0}^{\infty} e^{-nrT}$$

While it depends on the investment plan, we ignore the costs of acquiring initial capacity $k(0) = Ke^{-\delta T}$. The social problem then reduces to the choice of $q(t)$, K and T to maximize (1).

The choice of price can be de–coupled from investment decisions as long as the two are functionally separable. This condition holds here because the pricing problem depends only on capital stock $k(t) = Ke^{-\delta t}$ and not, say, on utilization rate. In that case, the solution has price set at marginal operating cost:

$$(3)\quad P(q^*(k)) = V_q(q^*(k),k)$$

along with the usual second–order conditions. This is just the solution to the *reverse* of the peak–load problem in which demand is fixed and capacity varies over time. Only current stock matters, and not peak capacity nor cycle length. Straightforward computation shows that, provided $V_{qk} < 0$, price *rises* over the cycle as scarce capital pushes up marginal (operating) cost. Consequently, as illustrated in Figure 1, price drops precipitously when each new cycle begins. It follows that price can change wildly with a small perturbation in demand or cost conditions.

Substituting optimal output into the firm's net operating revenue yields the quasi–rent $\pi(k) = R(q^*(k)) - V(q^*(k),k)$ which can be interpreted as a payment to capital holders. It is nonnegative since decreasing returns in the production technology ensures that marginal–cost pricing at least covers (variable) cost.[3] Moreover, the quasi–rent falls over the cycle which refutes a claim by Boiteux (1956b) that "sale at marginal cost involves deficits when the firm is over–equipped relative to demand, but it is profitable when the enterprise is very under–equipped." We will see that, while all periods contribute to construction expense, these payments may *not* be sufficient to defray the cost of the optimal investment project. Meanwhile we collect these observations in:

Proposition 1: Within an investment cycle, price is set at marginal variable cost, which falls over time, and the firm earns a nonnegative quasi–rent which also declines over the cycle.

The remaining problem is to choose K and T so as to:

$$\text{(4)}\quad \text{Maximize } W^* = \frac{1}{1-e^{-rT}}\left\{\int_0^T\left[S(q^*) - V(q^*,Ke^{-\delta t})\right]e^{-rt}dt - C(I)\right\}$$

where q^* is short–hand for $q^*(Ke^{-\delta t})$, similarly $S^* = S(q^*(Ke^{-\delta t}))$, and so forth. First–order

[3]This holds even if production exhibited increasing returns to scale. If the quasi–rent ever turned negative, the firm would do better by simply advancing the date of the next lumpy investment. In this way it realizes the profit of the next cycle sooner, and foregoes any depreciation that would occur while waiting.

conditions given the optimal price rule reduce to:

$$(5)\quad \frac{\partial W^*}{\partial K} = \frac{1}{1-e^{-rT}}\left[\int_0^T - V_k^* e^{-(\delta+r)t}dt - (1-e^{-\delta T})C'\right] = 0$$

$$(6)\quad \frac{\partial W^*}{\partial T} = \frac{-re^{-rT}}{(1-e^{-rT})^2}\left[\int_0^T (S^*-V^*)e^{-rt}dt - C\right] + \frac{1}{1-e^{-rT}}\left\{[S^*-V^*]_T e^{-rT} - \delta Ke^{-\delta T}C'\right\} = 0$$

The second–order conditions must also hold in addition to the sufficient conditions for an interior maximum, namely, $\partial W/\partial K > 0$ for all T when K = 0; likewise for changes in T. The rule for choosing peak capacity is straightforward: equate the marginal cost of construction to the marginal savings in terms of variable factors, both properly discounted. It is important to recognize that increasing returns in C is offset by depreciation and interest charges. Even with strong scale economies, investment will occur sporadically, and not at a single instant in time. Only in the extreme case when construction costs are invariant to the size of the project would all investment take place at the outset.

Recall that V is convex in (q,k) so that:

$$(7)\quad 0 = V(0,0) \geq V(q(t),k(t)) - V_q(q(t),k(t))q(t) - V_k(q(t),k(t))k(t)$$

where $k(t) = Ke^{-\delta t}$. A rearrangement gives:

$$(8)\quad \pi = Pq - V = V_q q - V \geq -V_k k.$$

Evaluating both sides of (8) at q^* and taking the present value yields:

$$(9)\quad \int_0^T (R^* - V^*)e^{-rt}dt \geq \int_0^T - V_k^* Ke^{-(\delta+r)t}dt = IC'(I)$$

The equality follows upon multiplying (5) by K. Now, the right–hand side of (9) will exceed construction expense whenever marginal cost exceeds average cost, or in other words, when $I^* > I^0$. However, this would be at variance with the maintained assumption of substantial economies in plant construction. We express our conclusions in a negative light:

Proposition 2: If optimal project size falls short of the minimum efficient scale, revenue will not cover variable cost plus investment expenses even when it is priced at marginal construction cost.

To generate other implications of these conditions, multiply (6) by $-(1-e^{-rT})/re^{-rT}$ to get:

$$\text{(10)} \quad \int_0^T (S^*-V^*)e^{-rt}dt - \left(\frac{1-e^{-rT}}{r}\right)[S^*-V^*]_T = C - \frac{1-e^{-rT}}{re^{-rT}}\,\delta K e^{-\delta T} C'$$

where $[\;]_T$ indicates the contents is evaluated at $t = T$. Collecting terms under the integral the left-hand side reduces to:[4]

$$\text{(11)} \quad \int_0^T [S^*-V^*]_T^t\, e^{-rt}dt = \int_0^T \left[-\int_t^T \frac{d(S^*-V^*)}{ds}\, ds\right] e^{-rt}dt$$

$$= \int_0^T\int_0^s \left[-\frac{d(S^*-V^*)}{ds}\, e^{-rt}\right] ds\, dt$$

$$= \int_0^T -\frac{d(S^*-V^*)}{ds}\, \frac{1-e^{-rs}}{r}\, ds$$

$$= -\frac{1}{r}[S^*-V^*]_0^T + \frac{1}{r}\int_0^T \frac{d(S^*-V^*)}{ds}\, e^{-rs}ds$$

$$= -\frac{1}{r}[S^*-V^*]_0^T + \frac{\delta}{r}\int_0^T V_k^*\, K e^{-(\delta+r)s}ds$$

$$= -\frac{1}{r}[S^*-V^*]_0^T - \frac{\delta}{r}C'I.$$

where the first equality follows by Calculus, the second by changing the order of integration, the third and fourth by integration, the fifth by totally differentiating the quasi–rent, and the last equality follows by substitution from (9). Finally, after inserting $Ke^{-\delta T} = K - I$, (11) can be rewritten as:

$$\text{(12)} \quad -[S-V]_0^T - \delta IC' = rC - \delta C'(e^{rt}-1)(K-I)$$

Rearranged and written out explicitly, (12) becomes:

$$\text{(13)} \quad [S(q(0)) - V(q(0),K) - \delta KC'(I)] - [S(q(T)) - V(q(T),Ke^{-\delta T}) - \delta(K-I)C'(I)e^{rT}] = rC(I).$$

The first term in square brackets is the surplus over variable costs and depreciation at the start

[4]The notation $[\;]_0^T$ refers to the difference between the value of the contents at $t = T$ and at $t = 0$.

of a cycle, where depreciation is valued at marginal cost of optimal investment. The second term is also the net surplus, but evaluated at the end of a cycle where depreciation is valued at the capitalized rate. The conclusion is succinctly expressed in:

Proposition 3:[5] The difference between the two net surpluses is equated to the competitive return on the construction outlay.

3. COST COVERAGE

Construction economies may render the firm insolvent when prices are set at marginal cost, but that does not mean it is unable to contribute some surplus to society. For this reason we turn to the problem of maximizing welfare while assuring the firm meets its obligations. When capital markets are perfect the profit constraint becomes:

$$(14) \quad \Pi = \frac{1}{1-e^{-rT}} \left\{ \int_0^T \left[P(q(t))q(t) - V(q(t),Ke^{-\delta t}) \right] e^{-rt} dt - C(K(1-e^{-\delta T})) \right\} \geq 0$$

The second–best problem amounts to maximizing (1) subject to (14). A simple way to do so is to replace total surplus S in the objective (1) with a weighted sum of surplus and revenue $(1-\rho)S + \rho R$. Here ρ is the "Ramsey–Boiteux number" that derives from the (nonnegative) multiplier λ on the firm's budget constraint they are related according to $\rho = \lambda/(1+\lambda)$. The number depends on K and T, but not on the date t. As ρ varies from 0 to 1, all solutions between the social optimum and the monopoly outcome are swept out. Barring discontinuities, an intermediate value will exist for which the firm will just break even (if that is at all possible).

The pricing subproblem yields the familiar inverse–elasticity rule:

$$(15) \quad \frac{P - V_q}{P} = \frac{\rho}{\varepsilon}$$

for all $t \in [0,T)$ where ε is the price elasticity of demand. The second–order conditions are stronger than when the budget constraint did not bind. Together with $V_{qk} < 0$, they once

[5](13) is a special case of Starrett's (second) rule of capacity pricing.

again imply that price rises over the cycle. The rate of change differs, however. For instance if demand is concave ($P'' < 0$) then price rises more slowly *ceteris paribus* than at the first–best, giving rise to a flatter price profile.

Besides current capital stock, output (and price) now depends on the Ramsey–Boiteux number: $q_\rho^*(Ke^{-\delta t})$. As in the static problem, Ramsey–Boiteux prices maximize surplus net of operating cost provided that the quasi–rent covers construction cost. Here the q_ρ^*'s are the solutions to a sequence of static budget–balancing problems which determine "Ramsey–Boiteux optimal payments to capital" to paraphrase Baumol, Panzar and Willig (1982). The properties of second–best prices are summarized in:

Proposition 4: At the second–best solution, the mark–up over marginal variable cost is inversely related to price elasticity. Also, price falls over time, as does the firm's quasi–rent.

The first–order conditions that characterize the second–best choice of K and T mimic (5) and (6):

$$(16)\quad \frac{\partial W^*}{\partial K} = \frac{1}{1-e^{-rT}}\left[\int_0^T - V_k^* e^{-(\delta+r)t}dt - (1-e^{-\delta T})C'(I)\right] = 0$$

$$(17)\quad \frac{\partial W^*}{\partial T} = \frac{-re^{-rT}}{(1-e^{-rT})}\int_0^T [(1-\rho)S^* + \rho R^* - V^* - C(I)]\, e^{-rt}dt + \frac{1}{1-e^{-rT}}\left\{[(1-\rho)S^* + \rho R^* - V^*]_T e^{-rT} - \delta K e^{-\delta T} C'(I)\right\} = 0$$

where W^* is present value of welfare when price is set according to (15). Denote the solutions to (16) and (17) as $K_\rho^*(T)$ and $T_\rho^*(K)$, respectively. For instance, given cycle length T, the first–best peak capacity is $K^*(T) = K_0^*(T)$ while the monopoly peak capacity is $\hat{K}(T) = K_1^*(T)$. Inspection of (16) reveals nothing new in the choice of peak capacity: the choice of peak capacity (given the cycle length) is not directly affected by the budget constraint, depending on ρ *only through output*. Just as in the static multiproduct problem, capital — or any factor of production for that matter — is chosen to minimize cost.

The innovation introduced by lumpy investment is found in the choice of cycle length. By lengthening the cycle, *both* price and investment are affected. Fix peak capacity and consider the effects of a rise in the cycle length. First of all additional investment is necessary to cover the additional physical depreciation. At the same time, however, this change forces price to continue to rise rather than jump downward at the end of the shorter cycle. Once investment does occur, more capacity will be on hand than before, justifying lower prices. A balance is struck in (17) between lower unit construction costs of the larger lumps and the delayed consumption.

To assess the effect of the budget constraint on the investment plan, we measure the change in peak capacity and cycle length as ρ varies. First observe that:

$$(18)\quad \frac{dK_\rho^*}{d\rho} = -\frac{\partial\, W^*}{\partial K \partial \rho} / \frac{\partial\, W^*}{\partial K^2}$$

so that the sign of $\partial\, W_\rho^*/\partial K \partial \rho$ is needed. Now,

$$(19)\quad \frac{\partial W_\rho^*}{\partial \rho} = \frac{1}{1-e^{-rT}} \int_0^T (R^* - S^*)e^{-rt}dt$$

and hence,

$$(20)\quad \frac{\partial\, W_\rho^*}{\partial K \partial \rho} = \frac{1}{1-e^{-rT}} \int_0^T q^* P'(q^*) \frac{dq^*}{dk}\, e^{-(\delta+r)t}dt.$$

It is easy to show that $q_\rho^*(k)$ is increasing in k, and so (20) is negative. Thus, $dK_\rho^*/\partial\rho < 0$ for each T. Similarly,

$$(21)\quad \frac{\partial\, W_\rho^*}{\partial T \partial \rho} = \frac{re^{-rT}}{(1-e^{-rT})} \int_0^T [S^* - R^*]_T^t\, e^{-rt}dt$$

which is positive since the term in square brackets is net consumer surplus which is positive and decreasing over time. Therefore, $dT_\rho^*(K)/d\rho > 0$.

In sum, as more weight is attached to profit and less to consumer surplus, second–best peak capacity (cycle length) falls (rises) for a given cycle length (peak capacity).

In order to perform comparative statics exercises it is crucial to know whether $K_\rho^*(T)$ and $T_\rho^*(K)$ are increasing or decreasing in T and K. Unfortunately, this is impossible to decide without further restrictions. Take the case of $K_\rho^*(T)$. An increase in the cycle length has two opposing effects. It increases the project size $I = K(1-e^{-\delta T})$ calling for a reduction in peak capacity to limit the additional construction costs, and simultaneously reduces base capacity $Ke^{-\delta T}$ which should be balanced with an increase in peak capacity. Formally, dK^*/dT and dT^*/dK have the same sign as the partial $\partial W^*/\partial K\partial T$ which is indeterminate. We refer to Case 1 when both $K^*_\rho(T)$ and $T^*_\rho(K)$ are increasing (i.e., $\partial^2 W^*_\rho/\partial K\partial T > 0$), and if both are decreasing (i.e., $\partial^2 W^*_\rho/\partial K\partial T < 0$), as Case 2. We conclude that:

Proposition 5: Upon imposing the budget constraint, the firm's investment plans change according to:

Case 1: it is not the case that $K^*_{\rho*} > K^*$ and $T^*_{\rho*} < T^*$

Case 2: $K^*_{\rho*} < K^*$ and $T^*_{\rho*} > T^*$

where $\rho^* \in [0,1]$ is the Ramsey–Boiteux number.

The first of the two possibilities is illustrated in Figure 2; the other case is just the "tilted" version of the same picture. In both cases the effect on the size of the project is ambiguous. This can be verified by performing the above exercise after replacing peak capacity with project size. This ambiguity makes it difficult to compare capital intensity under first– and second–best outcomes, unless one employs some notion of "average" capacity. A natural measure is the unweighted average of capital stock over the cycle:

$$(22) \quad A \equiv \frac{1}{T}\int_0^T Ke^{-\delta t}dt = \frac{K(1-e^{-\delta T})}{\delta T} = \frac{I}{\delta T}$$

It is an easy calculation to show that a fall in K or a rise in T will decrease average capacity. Other departures from the first best investment plan are possible as well, and therefore, a second–best outcome could have a larger average capacity than the first best. In fact, monopoly may be *more* capital intensive in this sense than either optima.

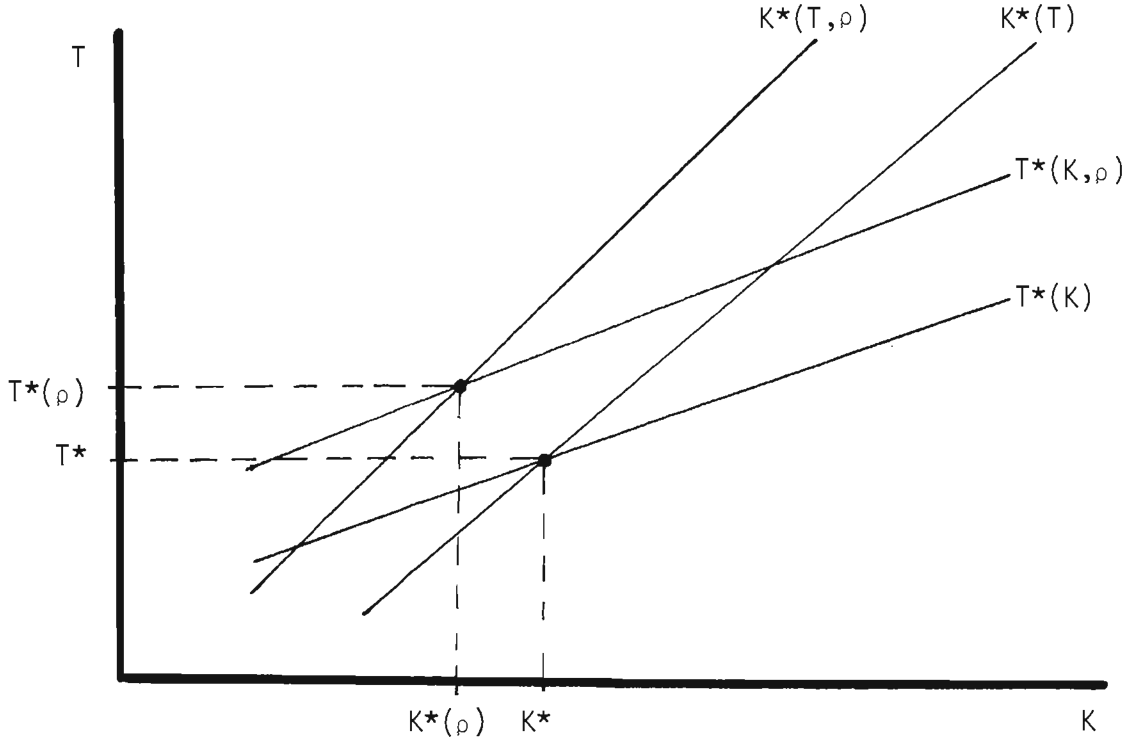

Figure 2: Effect of the Budget Constraint

More discouraging is the fact that the budget constraint has in indeterminate effect on prices. It is true that price unambiguously rises over the cycle as ρ goes from 0 to 1 *but only for a given investment plan.* It can happen that peak capacity falls and cycle length expands as the budget constraint is imposed thereby causing a drop in prices (at least over for the duration of the first cycle).

5. AN ALGEBRAIC EXAMPLE

The results produced up to this point can be illustrated with an algebraic example. It might also remove some of the ambiguity we have encountered along the way. The basic ingredients are the following demand and cost functions:

(23) $P(q) = q^{-1/\varepsilon}, \varepsilon > 1$

(24) $V(q,k) = q^{\alpha}k^{-\beta}, \ \alpha > 0, \beta > 0$

(25) $C(I) = I^{1/\gamma}, \gamma > 1$

Demand exhibits constant elasticity in which case:

(26) $S(q) = [\varepsilon/(\varepsilon-1)]q^{-1/\varepsilon}$

Operating costs can be viewed as generated by a generalized Cobb–Douglass production function where the degree of scale economies is given by $\mu = (1+\beta)/\alpha \leq 1$, i.e., non–increasing returns to scale. Scale economies in construction is measured by $\gamma > 1$ which is consistent with increasing returns to scale. In principle it is possible to derive the various price–investment policies and examine how they vary in terms of the parameters of demand (i.e., ε), operating cost (i.e., μ, δ, r) and especially construction cost (i.e., γ) and the social objective (i.e., ρ).

The mark–up over short–run marginal cost in (15) becomes:

(27) $$q^*(k) = \left[\frac{\alpha\varepsilon}{\varepsilon-\rho} k^{-\beta}\right]^{\varepsilon/(\varepsilon-\alpha\varepsilon-1)}$$

which is increasing in the stock of capital. Upon substituting $k = Ke^{-\delta t}$ we see immediately that price rises at the constant rate:

(28) $$\dot{p}^*/p^* = -\beta\delta/(\varepsilon-\alpha\varepsilon-1) > 0$$

which, incidentally, is independent of the budget constraint. Plugging output into the expression for welfare yields:

(29) $$W^* \equiv (1-\rho)S^* + \rho R^* - V^* = ak^b$$

where we have defined:

(30) $$a \equiv \left[\frac{\alpha\varepsilon}{\varepsilon-\rho}\right]^{\alpha\varepsilon/(\varepsilon-\alpha\varepsilon-1)} \left[\frac{\alpha\varepsilon-\varepsilon+1}{\varepsilon-1}\right]$$

(31) $$b \equiv \frac{-\beta(\varepsilon-1)}{\varepsilon-\alpha\varepsilon-1}$$

Under the assumptions, $a > 0$ and $0 < b < 1$. Note that pricing is affected by ρ only through the parameter a. A simple calculation demonstrates that the firm's quasi–rent takes a functional form similar to output:

(32) $$\pi^* \equiv R^* - V^* = Ak^b$$

where

(33) $$A = \left[\frac{\alpha\varepsilon}{\varepsilon-\rho}\right]^{(\varepsilon-1)/(\varepsilon-\alpha\varepsilon-1)} \left[\frac{\alpha\varepsilon-\varepsilon+\rho}{\alpha\varepsilon}\right] > 0$$

so that it is strictly positive and declining throughout the cycle.

Substituting for the expressions in (16) and (17) and rearranging to isolate K, we have:

$$(34)\quad K^{b-1/\gamma} = \frac{(\delta b+r)\ (1-e^{-\delta T})^{1/\gamma}}{ab\gamma[1-e^{-(\delta b+r)T}]}$$

$$(35)\quad K^{b-1/\gamma} = \frac{(\delta/\gamma)e^{-\delta T}(1-e^{-\delta T})^{1/\gamma-1}(1-e^{-rT}) - re^{-rT}(1-e^{-\delta T})^{1/\gamma}}{a(1-e^{-rT})e^{-(\delta b+r)T} - are^{-rT}[1-e^{-(\delta b+r)T}]/(\delta b+r)}$$

Equating the two yields a condition that implicitly determines T^*. Observe that whatever the solution, it will not depend on a, and hence, it will be *independent* of ρ. Thus, a profit–maximizing monopolist and a subsidized public enterprise alike would add lumps at intervals of the *same* length. From (27) and (32) we see that output and quasi–rent will move in same direction as K^*. Other less trivial comparative statics results are rather messy. However, if we make the not unreasonable assumption that $\delta = r$, this equation reduces to:

$$(36)\quad (1+b-b\gamma)e^{(1+b)rT} - (1+b)e^{rT} + b\gamma = 0$$

There is at least one, albeit uninteresting, solution to (36): T = 0. A sufficient condition for a strictly positive solution is that $1 + b - b\gamma > 0$. Immediately we have that T^* is independent of r and δ since neither appear in the coefficients of the polynomial.

Returning to (34) it is now easy to discern the effect of the constraint on peak capacity. Differentiating we have:

$$(37)\quad \frac{dK^*}{d\rho} = \frac{\gamma}{b\gamma-1} K^{(\gamma-b\gamma+1)/\gamma}\left\{\frac{-(\delta b+r)(1-e^{-\delta T})^{1/\gamma}}{a^2b\gamma[1-e^{-(\delta b+r)T}]}\right\}\frac{da}{d\rho}$$

which has the same sign as $b\gamma - 1$ since:

$$(38)\quad \frac{da}{d\rho} = \frac{-\alpha\varepsilon}{(\varepsilon-1)(\varepsilon-\rho)}\left[\frac{\alpha\varepsilon}{\varepsilon-\rho}\right]^{\alpha\varepsilon/(\varepsilon-\alpha\varepsilon-1)} < 0$$

in which case:

$$(39)\quad \frac{dK^*}{d\rho} \gtrless 0 \iff b\gamma - 1 \gtrless 0$$

Some tedious algebra establishes that $b\gamma \gtrless 1$ according to whether:

$$(40)\quad \alpha\frac{\varepsilon}{\varepsilon-1} - \beta\gamma \lesseqgtr 1$$

Recall that decreasing returns in production requires that $\alpha - \beta > 1$ so that the left side of (40) is greater than the right side provided $\varepsilon/(\varepsilon-1) > \gamma$, or in words, that demand is not too elastic nor returns to scale in construction too strong. Alternatively, an α large relative to β — that is, strong decreasing returns to scale in production — works in this same direction. In these cases, we conclude that $dK^*/d\rho < 0$. This is as expected since it says, relative to the first–best outcome, peak capacity is lower at the constrained solution. Likewise, the size of both the investment project and the average capacity of the firm falls because the cycle length does not change. All levels would be higher at the second–best than at the monopoly solution, in which case a monopolist undertakes projects that are too small compared with either the constrained or unconstrained optimum. It should be emphasized that demand and cost conditions that are entirely plausible can exactly reverse this conclusion.

Of the remaining comparative statics exercises we could perform the role of construction economies is most important. To assess the relation between construction economies and cycle length, totally differentiate (36) with respect to γ to find:

$$(41) \quad \frac{dT^*}{d\gamma} \gtrless 0 \iff b\gamma - 1 \gtrless 1 - e^{-brT^*}$$

which closely resembles (39). Since the right side is strictly positive, the conditions mentioned above that caused peak capacity to rise as we approached the first best (i.e., $b\gamma - 1 < 0$) would *a fortiori* reduce cycle length as increasing returns in construction became more pronounced. This is a odd since we would expect the increased efficiency of construction would induce a shift toward capital and away from other factors. Unfortunately we cannot complete the picture because little can be said about the relation between peak capacity and construction economies.

6. DECENTRALIZATION

A natural question at this stage is how to implement the second–best solution in practice. Implicit in the above discussion was the presence of some central authority that could intervene and impose the constrained solution on the enterprise. In reality legal and informational restrictions and administrative costs limit the ability of regulatory agencies to control every aspect of the firm's behavior. Furthermore, it is tempting to discover some way to harness the well–known powers of the decentralized market to seek out efficient outcomes.

For these reasons we now evaluate the performance of several alternative policies of decentralized control.

To begin with, investment planning is delegated to the firm while price control is retained by the central authority. A second approach reverses the roles of the two so that the firm has control over investment. Not surprisingly neither policy achieves the second best. Also, a complete comparison of the two alternatives hinges on the specific demand and cost conditions. We are inclined, nonetheless, to give price control the advantage over investment control on the grounds that at least it results in welfare–optimal choice of peak capacity. Finally we informally discuss the use of entry promotion as an alternative policy.

6.1 Price Control

As a starting point suppose that the firm is allowed to choose investment so as to maximize its profit. If prices are set according to the pricing rule given in (3) will the outcome be efficient? The firm's problem is then to choose K and T so as to:

$$(42) \quad \text{Maximize } \bar{\Pi} = \frac{1}{1-e^{-rt}} \left[\int_0^T (R^* - V^*) e^{-rt} dt - C(I) \right]$$

For a given cycle length T, profit–maximizing peak capacity $\bar{K}(T)$ satisfies:

$$(43) \quad \frac{\partial \bar{\Pi}}{\partial \bar{K}} = \frac{1}{1-e^{-rT}} \left\{ \int_0^T \left[(R_q^* - V_q^*) \frac{dq^*}{dk} - V_k^* \right] e^{-(\delta+r)t} dt - (1-e^{-\delta T}) C'(I) \right\} = 0$$

Provided the firm takes price as given, $dq^*/dk = 0$, in which case (42) is the same as (5) and hence $\bar{K}(T)$ and $K^*(T)$ coincide. In order to compare $\bar{T}(K)$ with $T^*(K)$, we attempt to sign the discrepancy between their respective first–order conditions. This is done in the Appendix. The effect of price control on investment again depends upon whether K and T are directly or indirectly related. In both cases $\bar{T} > T^*$; only the ordering of $\bar{K}$ and K^* is open to question. This is shown in Figure 3 for the first case and stated formally in:

Proposition 6: Delegating the size and frequency of investment to the firm departs from the first–best outcome according to:

Case 1: $\bar{K} > K^*$ and $\bar{T} > T^*$

Case 2: $\bar{K} < K^*$ and $\bar{T} > T^*$ so that $\bar{A} < A^*$

Incidentally, the orderings of peak capacity and cycle length are unchanged if marginal–cost pricing is replaced by the Ramsey–Boiteux inverse–elasticity rule. Hence neither the first–best nor the second–best outcomes can be decentralized through price control alone which leads us to consider a second form of decentralization.

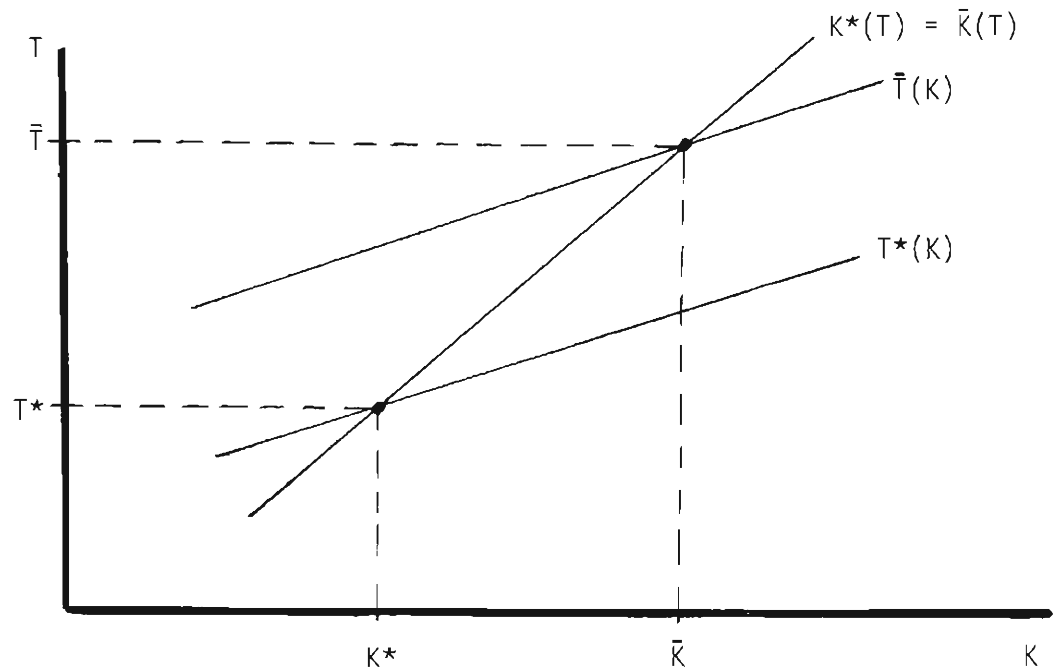

Figure 3: Effect of Price Control

6.2 Investment Control

Suppose the authority attempts to induce efficient allocation by controlling the size and frequency of plant investments. Once again, the first–best outcome is unattainable. Verification proceeds along the same lines as in the previous section except here the reference point will be the monopoly outcome, and not the welfare optimum. The social objective is:

$$(44) \quad \text{Maximize } \bar{\bar{W}} = \frac{1}{1-e^{-rT}} \left\{ \int_0^T [S(\hat{q}) - V(\hat{q}, Ke^{-\delta t})]e^{-rt}dt - C(I) \right\}$$

where $\hat{q}(k)$ is monopoly output solving $P + qP' = V_q$ for each k. Peak capacity and cycle length are chosen to maximize (44). In the Appendix we demonstrate that when peak capacities and cycle lengths are directly related, it will *not* be the case that investment control leads to a lower peak capacity and a longer cycle length. This is illustrated in Figure 4 once again just for Case 1. When they are indirectly related, peak capacity is higher and cycle length shorter after investment control is imposed, in which case average capital stock will increase as well.

Proposition 7: The effect of investment regulation relative to the monopoly outcome is given by:

Case 1: it is not the case that $\bar{\bar{K}} < \hat{K}$ and $\bar{\bar{T}} > \hat{T}$

Case 2: $\bar{\bar{K}} > \hat{K}$ and $\bar{\bar{T}} < \hat{T}$ so that $\bar{\bar{A}} > \hat{A}$.

One instance of investment control takes the form of the infamous "regulatory lag". In the present context, it amounts to an extension of the investment cycle beyond the monopoly level $\hat{T}$. We have here one more instance in which lagged effect of competition can have a beneficial effect.

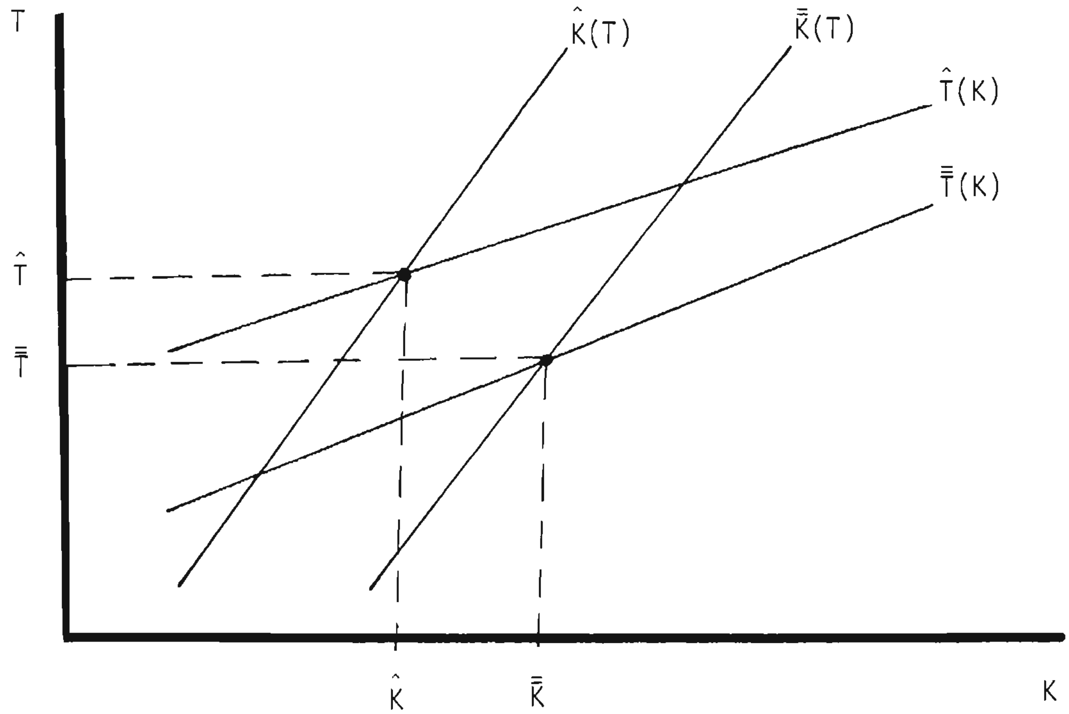

Figure 4: Effect of Investment Control

6.3 Some Remarks on Entry Control

Up to this point the firm has operated free of any sort of competition in its product market. Protection from potential competitors is often justified as necessary to ensure the firm's viability after it is forced to relinquish complete or partial control of its pricing and investment decisions. Alternatively, if the market is deemed to be a natural monopoly, then any additional firms necessarily raise the total cost of production. In spite of its cost effectiveness, however, Baumol, Panzar and Willig (1982) found that a natural monopoly — even a well–behaved one — may be vulnerable to entry threats. Taken too far, of course, artificial entry barriers will foster monopoly excesses that support abnormal profits and eats into consumer surplus.

Sound judgment regarding entry control demands scrutiny of two factors. On the one hand, one needs to know the cost–efficient industry structure. Specifically, is the cost of producing the good subadditive, in which case it should be carried on by a single firm? On the other hand, the costs associated with setting up and shutting down the firm govern the ease of entry into a market, and consequently the credibility of an entrant's threat. Baumol et. al.

have demonstrated that the cost conditions and the ease of entry may work in the opposite direction of efficiency. We will argue that, in the case of lumpy investment, the act of inducing the correct level entry will be especially delicate when investment is lumpy.

The situation described in our model is pre–disposed to natural monopoly if only because of an intertemporal cost complementarity that arises. Durable, imperfectly irreversible plant and equipment contribute to production at many different points in time, and it is the public good aspect of such sharing of facilities that is a prime source of subadditivity.

For centuries economists have pointed to fixed costs as the source of cost advantage of monopoly production. As proposed in our model, depreciation may act like a fixed cost. Between investment episodes, capital undergoes physical deterioration at a constant geometric rate, which makes it independent of intensity of use.

One would expect the presence of scale economies in plant construction would be sufficient for cost subadditivity. In fact, they have little bearing on the natural monopoly status of the firm, even when there are fixed costs associated with adding a lump of capacity. This can be seen by starting with a stationary investment plan given by K^* and T^* and show that two firms can produce any output path $q(\cdot)$ cheaper than a single firm. Let each of two firms alternate adding a lump of size I^* at intervals of length $2T^*$ with one firm starting at $t = 0$ and the other delaying until $t = T^*$. As a consequence of geometric depreciation, the combined capital stock of the two firms exactly equals that of a single firm since it is the simple sum of each capital stock: $k^*(t) = k_1(t) + k_2(t) = K^* e^{-\delta t}$ for all $t \in [nT^*,(n+1)T^*)$ $n = 0,1,\ldots$ (Only if the two firms invested simultaneously would a single firm register lower unit construction costs.) Moreover, the operating cost of the two firms will be no higher than the monopolist since by convexity of V and $V(0,0) = 0$, we have:

(45) $V(\lambda q^*,\lambda k^*) \leq \lambda V(q^*,k^*)$

for any $\lambda \leq 1$. Taking $\lambda_i(t) = k_i(t)/k^*(t)$ we conclude:

(46) $$\sum_1^2 V(\lambda_i q^*(t),k_i^*(t)) \leq V(q^*(t),k^*(t))$$

which shows that the cost is no greater when total output is divided between the two firms in proportion to their current capital stock. Since this will not be the optimal allocation of output, costs could be reduced even further.

We draw the tentative conclusion that the firm in our model will tend to be a natural monopoly, but not because of the lumpiness of the investment, rather as a result of its durability and irreversibility.

Turning to the issue of entry, we are especially concerned with situations in which entrants are unprofitable even though efficiency dictates more than one firm.

The need to sink costs is one of the clearest sources of entry barriers. It is also a particularly stubborn form of entry barrier since it is based in part on technology. Sunk investment would occur in our set-up to the extent that construction expenses could not be recovered at all or in part once they were incurred. Whether this is an accurate depiction will depend on the context. Suffice it to say that the best examples of lumpy investment tend to be virtually immobile and to a great extent customized to their intended function, both qualities that contribute to sunkness.

A formal analysis of sustainability requires us to determine whether an entrant could produce some subset of dated commodities and turn a profit at the prevailing prices. We do not answer this question but instead indicate how it can be made substantially easier. Observe first that, since demands are independent over time, we only need to check whether the incumbent earns a profit net of operating cost at any single moment. Proper measurement of operating cost, and in particular, depreciation expense is needed to do this correctly. Since in our case there is no investment except at distant points in time, economic depreciation is simply the sum of the quasi-rent and the competitive return on the firm's assets. The second component would be zero, of course, if all investment was entirely sunk upon installation. A depreciation rule that deviated at all from economic depreciation — as they all do — would induce the wrong amount of entry into this market. The measurement problems that must be surmounted to design the correct depreciation policy is a formidable challenge for even the most enlightened regulator.

7. CONCLUSIONS

We have modified the standard pricing and investment rules for a public enterprise to accommodate discrete increments to capital stock that are warranted by construction economies. Before ending we want to argue that the lessons learned in the process extend well beyond the limited applications entertained by Boiteux and his colleagues at Electricite d'France.

First of all many other public utilities are likely to exhibit significant construction economies, particularly those which involve a network (e.g., laying a pipeline, or the launch of a satellite) or which constitute part of the economy's infrastructure (e.g., building a new bridge or boring a tunnel). The same holds for unregulated, for–profit industries, especially when it is recognized that the phenomenon of lumpiness is not confined solely to capital investment decisions. As is well know, discontinuous jumps are an integral feature of optimal policies of inventory management and machine replacement. Also firms frequently drop a current production technique in favor of an alternative, or to switch from one supplier to a "second source". Finally, impulses may occur as well on the revenue side when a firm abruptly re–positions its product in the space of characteristics.

It is important to remember that in all of these cases the firm undergoes a discontinuity out of its free choice. Just as smooth accumulation was possible in our model, a gradual transition from one production technique or product design to another is entirely feasible. Firms willingly accept some instability in price, or quantity or other variables in exchange for the lower costs realized by "lumps." No doubt the firm's customers and suppliers may feel differently. Agents will typically be averse to such instability Especially when jumps incur adjustment costs. We can expect them to seek out ways to smooth out the discontinuities in their environment. For instance, in the electric power industry witness the practice of capacity sharing in nuclear powerplant projects and in power pooling arrangements.

In sum, lumpiness not only compels us to re–consider marginal rules developed over the years to guide pricing and investment decisions, but also challenges us to evaluate our understanding of the ability of market institutions and organizational arrangements to mitigate the disruptions caused by discontinuous behavior.

References

W. Baumol, J. Panzar, R. Willig (1982), Contestable Markets and the Theory of Industrial Structure, New York: Harcourt, Brace.

M. Boiteux (1949), "Peak load pricing," translated in the Journal of Business, April 1960.

M. Boiteux (1956a), "On the management of public monopolies subject to budgetary constraint," translated in Journal of Economic Theory, September 1971.

M. Boiteux (1956b), "Marginal cost pricing," in Revue Francaise del'Energie, translated in Marginal Cost Pricing in Practice (1964), J. Nelson (ed.), New York: Prentice-Hall.

W. Brock and W. Dechert (1985), "Dynamic Ramsey pricing, "International Economic Review.

H. Chenery (1952), "Overcapacity and the acceleration principle," Econometrica, 20:1, January, 1–28.

H. Mohring (1970), "The peak load problem with increasing returns and pricing constraints," American Economic Review, March, 693–705.

D. Starrett (1978), "Marginal cost pricing of recursive lumpy investments," Review of Economic Studies, 215–227.

O. Williamson (1966), "Peak-load pricing and optimal capacity under indivisibility constraints," American Economic Review, 56, September, 810–27.

Appendix

Derivation of Proposition 6:

For given K, the profit–maximizing cycle length, $\overline{T}(K)$, satisfies:

$$\text{(A1)}\quad \frac{\partial\overline{\Pi}}{\partial T} = \frac{-re^{-rT}}{(1-e^{-rT})}\left\{\int_0^T (R^*-V^*)e^{-rt}dt - C\right\} + \frac{1}{1-e^{-rT}}\left\{[R^*-V^*]_T\, e^{-rT} - Ke^{-\delta T}C'\right\} = 0$$

Combined with (6), we have:

$$\text{(A2)}\quad \frac{\partial\overline{\Pi}}{\partial T} - \frac{\partial\Pi^*}{\partial T} = \frac{-re^{-rT}}{(1-e^{-rT})}\int_0^T (S^*-R^*)e^{-rt}dt + \frac{1}{1-e^{-rT}}[S^*-R^*]_T e^{-rT}$$

and this, in turn, is equal to:

$$\text{(A3)}\quad \frac{-re^{-rT}}{(1-e^{-rT})}\int_0^T [S^*-R^*]_T^t\, e^{-rt}dt$$

which is negative because

$$\text{(A4)}\quad [S^*-R^*]_T^t = S(q^*(t)) - R(q^*(t)) - [S(q^*(T)) - R(q^*(T))] > 0$$

since S^*-R^* is just consumer surplus and q^* decreases over time. Hence, the left hand side of (16) is negative when evaluated at $\overline{T}(K)$ so that $\overline{T}(K) > T^*(K)$.

Derivation of Proposition 7:

For given T, the best choice of K satisfies:

$$\text{(A5)}\quad \frac{\partial\overline{\overline{W}}}{\partial K} = \frac{1}{1-e^{-rT}}\left\{\int_0^T [P(\hat{q}) - V_q(\hat{q},Ke^{-\delta t}))\frac{d\hat{q}}{dk} - V_q(\hat{q},Ke^{-\delta t})]e^{-(\delta+r)t}dt - (1-e^{-rT})C'(I)\right\} = 0$$

Evaluating (15) at $\rho = 1$, (A5) can be rewritten as:

$$\text{(A6)}\quad \frac{\partial\overline{\overline{W}}}{\partial K} = \frac{\partial\hat{W}}{\partial K} - \frac{1}{1-e^{-rT}}\int_0^T \hat{q}P'(\hat{q})\frac{d\hat{q}}{dk}e^{-(\delta+r)t}dt.$$

The second term is negative since $d\hat{q}/dk$ is positive. Therefore, $\partial\overline{\overline{W}}/\partial K > \partial\hat{W}/\partial K$. Hence, $\overline{\overline{K}}(T) > \hat{K}(T)$. Finally, for given K, we have:

$$\text{(A7)}\quad \frac{\partial \bar{\bar{W}}}{\partial T} = \frac{-re^{-rT}}{(1-e^{-rT})}\int_0^T [\hat{S} - \hat{V}]_T^t e^{-rt}dt - \frac{\delta K e^{-\delta T} C'(I)}{1-e^{-rT}}$$

$$= \frac{\partial \hat{W}}{\partial T} - \frac{re^{-rT}}{(1-e^{-rT})}\int_0^T [\hat{S} - \hat{R}]_T^t e^{-rt}dt$$

the equality following upon evaluating (16) at $\rho = 1$. Since the integrand in (A8) is positive by the previous argument, $\partial \bar{\bar{W}}/\partial T < \partial \hat{W}/\partial T$, and so $\bar{\bar{T}}(K) < \hat{T}(K)$.