Irreversible Investment, Capacity Choice, and the Value of the Firm

By Robert S. Pindyck*

Most investment expenditures are irreversible. The firm invests until the full value of an incremental unit of capacity equals its full cost. The former includes the value of the option to not utilize the unit; the latter includes the opportunity cost of sinking resources in the unit. We examine implications for capacity choice, utilization, firm value, and long-run marginal cost. With irreversibility, capacity is smaller, and firm value can be largely attributable to growth possibilities.

Most major investment expenditures are at least partly irreversible: the firm cannot disinvest, so the expenditures are sunk costs. Irreversibility usually arises because capital is industry- or firm-specific, that is, it cannot be used in a different industry or by a different firm. A steel plant, for example, is industry-specific. It can only be used to produce steel, so if the demand for steel falls, the market value of the plant will fall. Although the plant could be sold to another steel company, there is likely to be little gain from doing so, so the investment in the plant must be viewed as a sunk cost. As another example, most investments in marketing and advertising are firm-specific, and so are likewise sunk costs.1

The irreversibility of investment has been neglected since the work of Kenneth Arrow (1968), despite its implications for spending decisions, capacity choice, and the value of the firm. When investment is irreversible and future demand or cost conditions are uncertain, an investment expenditure involves the exercising, or "killing," of an option—the option to productively invest at any time in the future. One gives up the possibility of waiting for new information that might affect the desirability or timing of the expenditure; one cannot disinvest should market conditions change adversely. This lost option value must be included as part of the cost of the investment. As a result, the Net Present Value (NPV) rule "Invest when the value of a unit of capital is at least as large as the purchase and installation cost of the unit" is not valid. Instead the value of the unit must exceed the purchase and installation cost, by an amount equal to the value of keeping the firm's option to invest these resources elsewhere alive—an opportunity cost of investing.

This aspect of investment has been explored in an emerging literature, and most notably by Robert McDonald and Daniel Siegel (1986). They show that with even moderate levels of uncertainty, the value of this opportunity cost can be large, and investment rules that ignore it will be grossly in error. Their calculations, and those in related papers by Michael Brennan and Eduardo Schwartz (1985) and Saman Majd and Robert Pindyck (1987), show that in many cases projects should be undertaken only when their present value is at least double their direct cost.2

---

1Partial irreversibility also results from the "lemons" problem. Office equipment, cars, trucks, and computers are not industry-specific, but have resale value well below their purchase cost, even if new.

2Other examples are Ben Bernanke, 1983; Alex Cukierman, 1980; Carliss Baldwin, 1982; and Jeffrey Mackie-Mason, 1988. In the papers by Bernanke and

*Massachusetts Institute of Technology, Cambridge, MA 02139. This is a revised version of my 1986 NBER Working Paper, and is based on research done while I was Visiting Professor of Economics and Fellow, Institute of Advanced Studies, at Tel-Aviv University. I am grateful to the National Science Foundation for research support under grant no. SES-8318990, and to Giuseppe Bertola, Olivier Blanchard, Hua He, Saman Majd, Robert McDonald, Julio Rotemberg, Lawrence Summers, and two anonymous referees for helpful discussions and comments. Any errors are mine.

969
The existing literature has been concerned with investment decisions involving discrete projects, for example, whether to build a factory. This paper examines capacity choice and expansion, for example, how large a factory to build, and when to expand it. In particular, I focus on the marginal investment decision. This provides a simple and intuitively appealing solution to the optimal capacity problem. It also yields insight into the sources of the firm's value, and clarifies the measurement of long-run marginal cost.  

A firm's capacity choice is optimal when the present value of the expected cash flow from a marginal unit of capacity just equals the total cost of that unit. This total cost includes the purchase and installation cost, plus the opportunity cost of exercising the option to buy the unit. An analysis of capacity choice therefore involves two steps. First, the value of an extra unit of capacity must be determined. Second, the value of the option to invest in this unit must be determined (it will depend in part on the value of the unit itself), together with the decision rule for exercising the option. In essence, this decision rule is the solution to the optimal capacity problem.

To determine the value of a marginal unit of capacity, we must account for the fact that if demand falls, the firm can choose not to utilize the unit. In effect, a unit of capacity gives the firm an infinite number of options to produce, one for every future time $t$, each with exercise price equal to production cost, and can be valued accordingly. As we will show, these "operating options" are worth more the more volatile is demand, just as a call option on a stock is worth more, the more volatile is the price of the stock. This suggests that the firm should hold more capacity when future demand is uncertain, but the opposite is true. The reason is that uncertainty also increases the value of the firm's investment options, and hence the opportunity cost of irreversibly investing. Although the value of a unit of capacity increases, this opportunity cost increases even more, so the net effect is to reduce the firm's optimal capacity.  

Note that a firm's market value has two components: the value of installed capacity (i.e., the value of the firm's options to utilize some or all of this capacity over time), and the value of the firm's options to add more capacity later. As we will see, numerical simulations suggest that for many firms, "growth options" should account for a substantial fraction of market value, and the more volatile is demand, the larger is this fraction.

Options to productively invest are important assets, which firms hold even if they are price-takers in product and input markets. How do they arise? In some cases it is the result of a patent on a production technology, or ownership of land or natural resources. More generally, a firm's managerial resources, reputation, market position, and possibly scale, all of which may have been built up over time, enable it to productively undertake investments that individuals or other firms cannot undertake.  

---

Cukierman, uncertainty over future market conditions is reduced as time passes, so firms have an incentive to delay investing when markets are volatile (for example, during recessions). In the other papers cited above and in the model I present here, future market conditions are always uncertain. Access to the investment opportunity is then analogous to holding a call option on a dividend-paying stock; an expenditure should be made only when the value of the resulting project exceeds its cost by a positive amount, and increased uncertainty raises the incentive to delay the expenditure. Thus the results are similar to those in Bernanke and Cukierman, but for different reasons. Option value also appears in the natural resource context: if future values of wilderness areas and parking lots are uncertain, it may be better to wait before irreversibly paving a wilderness area. See, for example, Claude Henry, 1974.

The analysis in this paper is closely related to that in Giuseppe Bertola, 1987, who independently developed a model of capacity choice similar to mine, using a different solution approach.

---

1 In Andrew Abel, 1983, and Richard Hartman, 1972, uncertainty over future prices leads to an increase in the firm's optimal capital stock when the production function is linear homogeneous. The reason is that the marginal revenue product of capital is a convex function of price, so that as in my model, a marginal unit of capital is worth more when price is stochastic. However, in Abel and Hartman investment is reversible, so the opportunity cost of investing is zero.

2 The importance of growth options and their implications for the firm's financial structure are discussed in...
Section I clarifies the nature of the firm’s options to invest and produce, and how they affect its choice of capacity and its market value. Section II then solves the capacity choice problem in the context of a specific model. As the model is developed, a numerical example is used to show how the value of a marginal unit of capital, the opportunity cost of investing, and the firm’s optimal capacity depend on current demand and uncertainty over future demand. Sections III, IV, and V use the model to study the value of the firm, the behavior of capacity and capacity utilization over time, and implications for the measurement of marginal cost. Section VI concludes with caveats and limitations.

I. Optimal Incremental Investment Decisions

Consider a firm facing a demand curve that shifts over time stochastically, so that future demands are uncertain. Let \( \theta \) denote the demand shift parameter, with \( \partial P(Q, \theta)/\partial \theta > 0 \). Suppose the firm can install units of capital one at a time, at a sunk cost \( k \) per unit, whenever it wishes. Letting \( K \) be the amount of capital currently in place, we can write the value of the firm, \( W \), as the sum of two parts:

\[
W = V(K; \theta) + F(K; \theta) \tag{1}
\]

\( V(K; \theta) \) is the value of the firm’s capital in place, that is, the present value of the expected flow of profits that this capital will generate, given the current value of \( \theta \). \( F(K; \theta) \) is the value of the firm’s “growth options,” that is, given that the firm has capital \( K \) in place and given the current value of \( \theta \), \( F(K; \theta) \) is the present value of any additional profits that might result should the firm add more capital in the future, less the present value of the cost of that capital. Note that \( F(K; \theta) \) is greater than the present value of the expected flow of net profits from anticipated future investments, because the firm is not committed to any investment path.

Equation (1) is just an accounting identity, but we can use it to gain insight into the firm’s investment problem by noting that units of capital are installed sequentially. Assuming for now that the units are discrete, we can number them in the order they are installed. Suppose units 1 through \( n \) have been installed so far. Then, suppressing \( \theta \), we can rewrite (1) by summing the value of each installed unit and the values of the options to install further units:

\[
W = \Delta V(0) + \Delta V(1) + \Delta V(2) \cdots + \Delta V(n-1) + \Delta F(n) + \Delta F(n+1) + \cdots \tag{2}
\]

Here \( \Delta V(j) \) is the value of the \( j + 1 \)st unit of capital, that is, the present value of the expected flow of incremental profits generated by unit \( j + 1 \). Of course the firm need not always utilize this (or any other) unit of capital. It has the option to utilize it at every point during its lifetime, and \( \Delta V(j) \) is equal to the value of this option. Section II shows how \( \Delta V(j) \) can be calculated. The firm must decide whether to install additional capital. Given that \( n \) units are in place, \( \Delta F(n) \) is the value of the option to buy one more unit, that is, unit \( n + 1 \), at any time in the future. If the firm exercises this option, it pays \( k \) and receives an asset worth \( \Delta V(n) \). The firm also gives up \( \Delta F(n) \), because once exercised, the option is dead—whether or not the firm later buys more capital, it has now paid for unit \( n + 1 \), and cannot disinvest. Hence \( \Delta F(n) \) is also a cost of investing in this unit. The full cost of investing is thus \( k + \Delta F(n) \), and this must be compared to the benefit \( \Delta V(n) \).

Once the firm buys unit \( n + 1 \), it then faces another problem: at what point should it exercise its option, worth \( \Delta F(n + 1) \), to buy unit \( n + 2 \), which is worth \( \Delta V(n + 1) \)? And so on. These options must be exercised sequentially, so the total value of the firm’s

Stewart Myers, 1977. A complete model of industry evolution would also describe the competitive processes through which firms obtain these options. Such a model is beyond the scope of this paper.
options to grow is

\[ F(n) = \sum_{j=n}^{\infty} \Delta F(j). \]

Letting these incremental units become infinitesimally small, we can write equation (2) for the value of a firm with capital stock \( K \) as

\[ W = \int_{0}^{K} \Delta V(v; \theta) \, dv + \int_{K}^{\infty} \Delta F(v; \theta) \, dv. \]

The firm's optimal capital stock \( K^* \) is that which maximizes its net value, \( W - kK^* \). Using (3), this maximization gives the following optimality condition that must hold if the firm is investing:

\[ \Delta V(K^*; \theta) = k + \Delta F(K^*; \theta). \]

Thus the firm should invest until the value of a marginal unit of capital, \( \Delta V(K; \theta) \), is equal to its total cost: the purchase and installation cost, \( k \), plus the opportunity cost \( \Delta F(K; \theta) \) of irreversibly exercising the option to invest in the unit, rather than waiting and keeping the option alive.\(^6\)

The firm's investment problem can therefore be solved by calculating \( \Delta V(K; \theta) \) and \( \Delta F(K; \theta) \), and using (4) to determine the optimal capacity \( K^*(\theta) \). In the next section we carry out these steps for a specific model.

II. A Model of Capacity Choice

Consider a firm that faces the following demand function:

\[ P = \theta(t) - \gamma Q. \]

(The firm might be a price-taker, in which case \( \gamma = 0 \).) Here \( \theta(t) \) evolves according to the stochastic process

\[ d\theta = \alpha \theta dt + \sigma \theta dz, \]

where \( dz \) is the increment of a Weiner process, that is, \( dz = \epsilon(t)(dt)^{1/2} \), with \( \epsilon(t) \) a serially uncorrelated and normally distributed random variable. Equation (6) says that the current value of \( \theta \) (and thus the current demand function) is known to the firm, but future values of \( \theta \) are unknown, and are lognormally distributed with a variance that grows with the time horizon. Thus even though information arrives over time (the firm observes \( \theta \) changing), future demand is always uncertain.\(^7\)

The firm's cost and production constraints are as follows: (i) each unit of capital can be bought at a fixed price \( k \) per unit; (ii) each unit of capital provides the capacity to produce one unit of output per time period, so that \( Q \leq K \); and (iii) the firm has an operating cost \( C(Q) = c_1 Q + (1/2)c_2 Q^2 \). In general \( c_1 \) and/or \( c_2 \) can be zero, but if \( \gamma = 0 \) (so the firm is a price-taker), we require \( c_2 > 0 \) to bound the firm's size.

I assume that the firm starts with no capacity, so at \( t = 0 \) it must decide how much initial capacity to install. Later it might add more capacity, depending on how demand evolves. For simplicity I assume that new capacity can be installed instantly, and capi-

\(^6\)Note that \( \Delta V(K) \) is not the marginal value of capital, as the term is used in marginal \( q \) and related models of investment, such as the one in Abel and Olivier Blanchard, 1986. The marginal value of capital is the present value of the expected flow of profits throughout the future from whatever unit of capital is the marginal one, that is,

\[ E \int_{0}^{\infty} \left[ \frac{\partial \pi(K_t)}{\partial K_t} \right] e^{-\mu t} \, dt, \]

where \( \mu \) is the discount rate. The marginal value of capital thus depends on the firm's capital stock, \( K_t \), or its distribution, at every future \( t \), and so its calculation can be quite difficult. Note how this differs from \( \Delta V(K) \), the present value of the expected flow of incremental profits attributable to the \( K + 1 \)st unit of capital, which is independent of how much capital the firm has in the future.

\(^7\)Analytic solutions can be obtained for any demand function linear in \( \theta \) or a power function of \( \theta \); I use (5) for simplicity. Also, it is straightforward to also allow for uncertainty over future operating costs. The qualitative results would be the same.
tal in place does not depreciate. Finally, investment is completely irreversible—the firm cannot disinvest. This means there is an opportunity cost to investing. Adding a unit of capacity today precludes the possibility of waiting and instead adding the unit later or not at all.

I make one more assumption: stochastic changes in demand are spanned by existing assets, that is, there is an asset or dynamic portfolio of assets whose price is perfectly correlated with θ. (This is equivalent to saying that markets are sufficiently complete that the firm’s decision to invest or produce does not affect the opportunity set available to investors.) This assumption holds for most commodities, which are usually traded on both spot and futures markets, and for manufactured goods to the extent that prices are correlated with the values of shares or portfolios. However, in some cases this assumption will not hold, for example, a new product unrelated to any existing ones.

With the spanning assumption, we can determine the investment rule that maximizes the firm’s market value, and the investment problem reduces to one of contingent claim valuation. This provides insight, and avoids assumptions about risk preferences or discount rates. However, as shown in the Appendix, the problem can also be expressed in terms of dynamic programming. If spanning does not hold, dynamic programming can still be used to maximize the present value of the firm’s expected flow of profits, using an arbitrary discount rate. (Note that in such cases there is no theory for determining the discount rate; the CAPM, for example, would not hold.) ∆V(K) and ∆F(K) would satisfy differential equations nearly the same as those that will be derived below (but containing the discount rate), so the qualitative results will be the same.

Let x be the price of an asset or portfolio of assets perfectly correlated with θ, and denote by ρθx the correlation of θ with the market portfolio. Then x evolves according to

$$dx = μx \, dt + σx \, dz,$$

and by the CAPM, its expected return is $μ = r + φρθxσ$, where $φ$ is the market price of risk. I will assume that $φ$, the expected percentage rate of change of θ, is less than μ. (It will become clear later that if this were not the case, no firm in the industry would ever install any capacity. No matter what the current level of θ, firms would always be better off waiting and simply holding the option to install capacity in the future.) Denote the difference between μ and α by δ, that is, $δ = μ - α$.

For purposes of comparison, note that if future demand were certain ($σ = 0$), and if $α = 0$, the firm’s optimal initial capacity would be $K^∗(θ) = (θ - c_1 - rk)/(2γ + c_2)$, where r is the riskfree rate of interest. Equivalently, the firm should add capacity only if $θ(K) > (2γ + c_2)K + c_1 + rk$. We will see that uncertainty makes the optimal capacity smaller than this.

A. The Value of a Marginal Unit of Capacity

To solve the firm’s investment problem we first determine $∆V(K)$, the present value of the expected flow of profits from an incremental unit of capacity, given a current capacity K. Because the unit need not be utilized, the profit it generates at any future time t is a nonlinear function of θ, which is

---

8Relaxing these assumptions makes no qualitative difference in the results. In fact, allowing for lead times in the construction and installation of new capacity magnifies the effects of uncertainty, as shown in Majd and Pindyck, 1987.

9For an overview of contingent claims analysis and its applications, see John Cox and Mark Rubinstein, 1985; and Scott Mason and Robert Merton, 1985.

10The connection between the contingent claims approach that I use and dynamic programming is examined in detail by Bertola, 1987.

11If $α > 0$, the optimal initial capacity would be smaller than this. The reason is that the cost of investing is fixed but the payoff is growing, so there is a benefit to delay, and $∆F(K) > 0$ for all K.
stochastic:

\[ \Delta \pi_t(K) = \max[0, (\theta_t - (2\gamma + c_2)K - c_1)]. \]

Thus \( \Delta V(K) \) can be written as

\[ \Delta V(K) = \int_0^\infty \int_0^\infty \Delta \pi_t(K; \theta_t, t) \, d\theta_t e^{-\mu t} \, dt, \]

where \( f(\theta, t) \) is the density function for \( \theta \) at time \( t \), and \( \mu \) is the risk-adjusted discount rate. It is difficult, however, to evaluate (8) directly. In addition, the discount rate \( \mu \) might not be known.

Instead we obtain \( \Delta V(K) \) by solving the following equivalent problem: What is the value of a factory that produces 1 unit of output per period, with operating cost \((2\gamma + c_2)K + c_1\), which it sells in a perfectly competitive market at a price \( \theta \), and where the factory can be shut down (temporarily and costlessly) if \( \theta \) falls below the operating cost?\(^{12}\) The Appendix shows that the solution to this problem, obtained using contingent claim valuation methods or equivalently via dynamic programming, is:

\[ \Delta V(K) = \begin{cases} b_1 \theta \beta_1, & \theta \leq (2\gamma + c_2)K + c_1 \\ b_2 \theta \beta_2 + \theta / \delta - [(2\gamma + c_2)K + c_1] / r; & \theta \geq (2\gamma + c_2)K + c_1, \end{cases} \]

where

\[ \beta_1 = -\frac{(r - \delta - \alpha^2/2)}{\alpha^2} \]
\[ + \frac{1}{\alpha^2} \left[ (r - \delta - \alpha^2/2)^2 + 2r \sigma^2 \right]^{1/2} > 1 \]
\[ \beta_2 = -\frac{(r - \delta - \alpha^2/2)}{\alpha^2} \]
\[ - \frac{1}{\alpha^2} \left[ (r - \delta - \alpha^2/2)^2 + 2r \sigma^2 \right]^{1/2} < 0 \]
\[ b_1 = \frac{r - \beta_2 (r - \delta)}{r \delta (\beta_1 - \beta_2)} \times \]
\[ [(2\gamma + c_2)K + c_1]^{1 - \beta_1} > 0 \]
\[ b_2 = \frac{r - \beta_1 (r - \delta)}{r \delta (\beta_1 - \beta_2)} \times \]
\[ [(2\gamma + c_2)K + c_1]^{1 - \beta_2} > 0 \]

This solution for \( \Delta V(K) \) is interpreted as follows. When \( \theta < (2\gamma + c_2)K + c_1 \), the unit of capacity is not utilized. Then, \( b_1 \theta \beta_1 \) is the value of the firm’s option to utilize the unit in the future, should \( \theta \) increase. When \( \theta \geq (2\gamma + c_2)K + c_1 \), the unit is utilized. If, irrespective of changes in \( \theta \), the firm had no choice but to continue utilizing the unit throughout the future, the present value of the expected flow of profits would be given by \( \theta / \delta - [(2\gamma + c_2)K + c_1] / r \). (Costs are certain and so are discounted at the riskfree rate; future values of \( \theta \) are discounted at the risk-adjusted rate \( \mu \), but \( \theta \) is expected to grow at rate \( \alpha \), so the effective net discount rate is \( \mu - \alpha = \delta \).) However, should \( \theta \) fall, the firm can reduce output and not utilize this unit of capacity. The value of this option is \( b_2 \theta \beta_2 \).

A numerical example will help to illustrate the model. For this purpose I choose \( r = \delta = .05 \), \( k = 10 \), \( c_1 = 0 \), and either \( \gamma = .5 \) and

---

\(^{12}\) The valuation of a factory that can be temporarily shut down has been studied by Brennan and Schwartz, 1985; and McDonald and Siegel, 1985. Observe from equation (7) that the present value of an incremental profit at future time \( t \) is the value of a European call option, with expiration date \( t \) and exercise price \((2\gamma + c_2)K + c_1\), on a stock whose price is \( \theta \), paying a proportional dividend \( \delta \). This point was made by McDonald and Siegel, 1985. Thus \( \Delta V(K) \), the value of our “equivalent factory,” can be found by summing the values of these call options for every future \( t \). However this does not readily yield a closed form solution, and I use an approach similar to that of Brennan and Schwartz, 1985.
Figure 1 shows $\Delta V(K)$ as a function of $\theta$ for $K = 1$ and $\sigma = 0$, .2, and .4. Observe that $\Delta V(K)$ looks like the value of a call option —indeed it is the sum of an infinite number of European call options (see fn. 12). As with a call option, $\Delta V(K)$ is increasing with $\sigma$, and for $\sigma > 0$, $\Delta V(K) > \Delta V_0(K)$ because the firm need not utilize its capacity. As $\theta \to \infty$, $\Delta V(K) \to \Delta V_0(K)$; for $\theta$ very large relative to $K$, this unit of capacity will almost surely be continuously utilized over a long period of time.

Figure 2 shows $\Delta V(K)$ as a function of $K$ for $\theta = 2$, and $\sigma = .0, .1, .2, .4$. Because demand evolves stochastically, a marginal unit of capacity has some positive value no matter how large is the existing capital stock; there is always some chance that it will be utilized over any finite period of time. The greater is $\sigma$, the more slowly $\Delta V(K)$ declines with $K$. Also, the smaller is $K$, the more likely it is that the marginal unit will be utilized, and so the smaller is $\Delta V(K) - \Delta V_0(K)$. When $K = 0$, $\Delta V(0) = \Delta V_0(K)$; with $c_1 = 0$, the first marginal unit will always be utilized.

The fact that $\Delta V(K)$ is larger when $\sigma > 0$ might suggest that the firm should hold more
capacity, but the opposite is true. As shown below, uncertainty also makes the firm’s opportunity cost of exercising its option to invest in the marginal unit larger, and by an even greater amount.

B. The Decision to Invest in the Marginal Unit

Having valued the marginal unit of capacity, we can now value the firm’s option to invest in this unit, and determine the optimal decision rule for investing. In the Appendix it is shown that the value of the firm’s option to invest, $\Delta F(K)$, is

$\Delta F(K) = \begin{cases} a\theta^\beta; & \theta \leq \theta^*(K) \\ \Delta V(K) - k; & \theta \geq \theta^*(K) \end{cases}$

where

$\beta_1, \beta_2, \beta_3$ are given under equation (9) above, and $\theta^*(K)$ is the critical value of $\theta$ at or above which it is optimal to purchase the marginal unit of capacity, that is, the firm should purchase the unit if $\theta \geq \theta^*(K)$. The critical value $\theta^*(K)$ is in turn the solution to

$\beta_1 \theta_2 \beta_3 \beta_1 \beta_2 \beta_1 \beta_1 - \frac{\beta_1 - 1}{\beta_1 \beta_1 \beta_1} \theta^* - \frac{(2\gamma + c_2)K + c_1}{r} - k = 0$.

Equation (11) can be solved numerically for $\theta^*$, and equation (10) can then be used to calculate $\Delta F(K)$.

Recall our assumption that $\delta > 0$. The reader can verify that as $\delta \to 0$, $\theta^*(K) \to \infty$. Unless $\delta > 0$, the opportunity cost of investing in a unit of capacity always exceeds the benefit, and the firm will never install capacity. Thus if firms in an industry are invest-

$\Delta V(K)$ is a function of $\theta$, and if $\delta = 0$, $\theta$ is expected to grow at the risk-adjusted market rate. Since the option to invest is perpetual, there would be no gain from installing capacity now rather than later.
ing optimally and some positive amount of investment is taking place, we should observe $\delta > 0$.

As with a call option on a dividend-paying stock, both $\Delta F(K)$ and the critical value $\theta^*(K)$ increase as $\sigma$ increases. Figure 3 shows $\Delta F(K)$ as a function of $\theta$ for $K=1$ and $\sigma = 0$, .2, and .4. In each case $\theta^*$ is indicated by a “.” When $\sigma = 0$, $\theta^* = 1.5$, that is, the firm should increase capacity only if $\theta$ exceeds 1.5. For $\sigma = .2$ and .4, $\theta^*$ is 2.45 and 3.44, respectively. The opportunity cost of exercising the firm’s option to invest in additional capacity is $\Delta F(K)$, which increases with $\sigma$, so a higher $\sigma$ implies a higher critical value $\theta^*(K)$. Also, it is easily shown that $\theta^*(K)$ is monotonically increasing in $K$.

C. The Firm’s Optimal Capacity

The function $\theta^*(K)$ is the firm’s optimal investment rule; if $\theta$ and $K$ are such that $\theta > \theta^*(K)$, the firm should add capacity, increasing $K$ until $\theta^*$ rises to $\theta$. Equivalently we can substitute for $b_2(K)$ and rewrite equation (11) in terms of $\tilde{K}^*(\theta)$, the firm’s optimal capacity:

\[
\begin{align*}
(11') & \quad \frac{r - \beta_1(r - \delta)}{r \delta \beta_1} \theta^{\beta_1} [(2\gamma + c_2)K^* + c_1]^{1-\beta_2} \\
& - \frac{[(2\gamma + c_2)K^* + c_1]}{r} \\
& + \frac{\beta_1 - 1}{\delta \beta_1} \theta - k = 0.
\end{align*}
\]

Figure 4 shows $K^*(\theta)$ for $\sigma = 0$, .2, and .4. (For many industries .2 is a conservative value for $\sigma$—see fn. 14.) Observe that $K^*$ is much smaller when future demand is uncertain. For $\sigma = .4$, $\theta$ must be more than three times as large as when $\sigma = 0$ before any capacity is installed.

Another way to see how uncertainty over future demand affects the firm’s optimal capacity is by comparing $\Delta F(K)$, the value of the option to invest in a marginal unit, with $\Delta V(K) - k$, the net (of purchase cost) value of the unit. The optimal capacity $K^*(\theta)$ is
Figure 4. Optimal Capacity, $K^*$, as a Function of $\theta$

Figure 5. The Net Value of a Marginal Unit of Capacity, and the Value of the Option to Install the Marginal Unit
the maximum $K$ for which these two quantities are equal. Note from equation (10) that for $\theta \geq \theta^*$, or equivalently, $K \leq K^*$, exercising the option to invest maximizes its value, so that $\Delta F(K) = \Delta V(K) - k$, but for $K > K^*$, $\Delta F(K) > \Delta V(K) - k$, and the option to invest is worth more "alive" than "dead."

This is shown in Figure 5, which plots $\Delta F(K)$ and $\Delta V(K) - k$ as functions of $K$, for $\theta = 2$ and $\sigma = .2$. Recall that $\Delta V(K)$ is larger when future demand is uncertain. As the figure shows, if the opportunity cost of exercising the option to invest were ignored, that is, if the firm adds capacity until $\Delta V(K) - k = 0$, then capacity would be about 2.3 units (as opposed to 1.5 units when $\sigma = 0$). But at these capacity levels the opportunity cost of investing in a marginal unit exceeds the net value of the unit, so the value of the firm is not maximized. The optimal capacity is only $K^* = .67$, the largest $K$ for which $\Delta F(K) = \Delta V(K) - k$, and the solution to equation (11).

III. The Value of The Firm

As noted above, $K^*(\theta)$ maximizes the firm's market value, net of cash outlays for the purchase of capital. Recall that the firm's net value as a function of its capacity $K$ is given by:

\begin{equation}
\text{Net Value} = \int_0^K \Delta V(\nu) \, d\nu + \int_K^\infty \Delta F(\nu) \, d\nu - kK.
\end{equation}

Differentiating with respect to $K$ shows that this is maximized when $K = K^*$ such that $\Delta V(K^*) - \Delta F(K^*) - k = 0$.

The value of the firm's installed capacity, $V(K^*)$, is the first integral in equation (12). In Figure 5 it is the area under the curve $\Delta V(K) - k$ from $K = 0$ to $K^*$, plus the purchase cost $kK^*$. The value of the firm's growth options is the second integral, which in Figure 5 is the area under the curve $\Delta F(K)$ from $K = K^*$ to $\infty$. As the figure shows, growth options can account for a large portion of the firm's total value.

The sensitivity of firm value and its components to uncertainty over future demand can be seen from Table 1, which shows $K^*$, $V(K^*)$, $F(K^*)$, and total value for different values of $\sigma$ and $\theta$. When $\sigma = 0$, the value of the firm is only the value of its installed capacity. Whatever the value of $\theta$, the firm is worth more the more volatile is demand. A larger $\sigma$ implies a larger value for each unit of installed capacity, and a much larger value for the firm's options to expand. Also, the larger is $\sigma$, the larger is the fraction of firm value attributable to its growth options. When $\sigma = .2$ or more, more than half of the firm's value is $F(K^*)$, the value of its growth options. Even when $\sigma = .1$, $F(K^*)$ accounts for more than half of total value when $\theta$ is 1 or less. (When demand is currently small, it is the possibility of greater demand in the future that gives the firm much of its value.) And there is always a range of $\theta$ for which $K^*$ is zero, so that all of the firm's value is due to its growth options.

As mentioned earlier, a $\sigma$ of .2 or more would not be unusual. Thus an implication of the model is that for many firms, the fraction of market value attributable to the value of capital in place should be one-half or less. A second implication is that this fraction should be smaller the greater is the volatility of market demand. I have not tried to test either of these implications (valuing capital in place is difficult). However, calculations reported by Carl Kester (1984) are consistent with both of them. He estimated the value of capital in place for 15 firms in 5 industries by capitalizing the implied flows of anticipated earnings, and found that it is half or less of market value in the majority of cases. Furthermore, this fraction is only about 1/5 to 1/3 in industries where demand is more volatile (electronics, computers), but more than 1/2 in industries with less volatile demand (tires and rubber, food processing).

IV. The Dynamics of Capacity, Capacity Utilization, and Firm Value

If the firm begins with no capacity, it initially observes $\theta$ and installs a starting capacity $K^*(\theta)$. If $\theta$ then increases, it will
Table 1 — Value of Firm

\[ c_1 = c_2 = 0, \gamma = .5, r = \delta = .05, k = 10 \]

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>( \theta )</th>
<th>( K^* )</th>
<th>( V(K^*) )</th>
<th>( F(K^*) )</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>7.5</td>
<td>0</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.5</td>
<td>37.5</td>
<td>0</td>
<td>37.5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.5</td>
<td>87.5</td>
<td>0</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.5</td>
<td>157.5</td>
<td>0</td>
<td>157.5</td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>.5</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>1</td>
<td>0.23</td>
<td>4.3</td>
<td>3.4</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>33.1</td>
<td>9.1</td>
<td>42.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.82</td>
<td>80.4</td>
<td>20.3</td>
<td>100.7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.65</td>
<td>147.3</td>
<td>35.5</td>
<td>182.8</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>.5</td>
<td>0</td>
<td>0</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>1</td>
<td>0.04</td>
<td>0.8</td>
<td>13.5</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.67</td>
<td>24.0</td>
<td>49.2</td>
<td>73.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.37</td>
<td>67.1</td>
<td>94.6</td>
<td>161.7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.09</td>
<td>134.1</td>
<td>143.7</td>
<td>277.8</td>
<td></td>
</tr>
<tr>
<td>.4</td>
<td>.5</td>
<td>0</td>
<td>0</td>
<td>25.8</td>
<td>25.8</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>69.7</td>
<td>69.7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>5.9</td>
<td>182.6</td>
<td>188.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.64</td>
<td>36.2</td>
<td>307.2</td>
<td>343.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.22</td>
<td>91.3</td>
<td>427.5</td>
<td>518.8</td>
<td></td>
</tr>
</tbody>
</table>

expand capacity accordingly, and the value of the firm will rise. The value of its growth options will also rise, but will become a smaller fraction of total value (see Table 1). However, if \( \theta \) decreases, the firm will find itself holding more capacity than it would have chosen had the decrease been anticipated. The firm’s value will fall, and depending on how much \( \theta \) decreases, some of its capacity may become idle.

Because capital does not depreciate in this model, the firm’s capacity is nondecreasing, but will rise only periodically. The dynamics of capacity are characterized in Figure 6, which shows a sample path for \( \theta(t) \), and the corresponding behavior of \( K(t) \). (The duration of continuous upward movements in \( K(t) \) is exaggerated.) The firm begins at \( t_0 \) by installing \( K^*(\theta_0) = K_0^* \). Then \( \theta \) increases until it reaches a (temporary) maximum \( \theta_1 \) at \( t_1 \), and \( K \) is increased accordingly to \( K_1^* \). Here remains fixed until \( t_2 \), when \( \theta \) again reaches \( \theta_1 \). Afterward \( K \) is increased as \( \theta \) increases, until \( t_3 \) when \( \theta \) begins to decline from a new maximum, and \( K \) remains fixed at \( K_3^* \).

Thus an implication of the model is that investment occurs only in spurts, when demand is rising, and only when it is rising above historic levels.\(^{16}\) Firms usually increase capacity only periodically, and this is often attributed to the “lumpiness” of investment. But lumpiness is clearly not required for this behavior.

Let us now examine the firm’s capacity utilization. Clearly during periods of expansion, all capacity will be utilized. When demand falls, however, some capacity may go unutilized, but only if it falls far enough.

If the firm had unlimited capacity it would maximize current profits by setting output at \( Q^* = (\theta - c_1)/(2\gamma + c_2) \). However, \( K^*(\theta) < (\theta - c_1)/(2\gamma + c_2) \), and as shown in Section II, can be much less even for moderate values of \( \sigma \). Thus for \( \theta \) in the range \( \theta(K) = (2\gamma + c_2)K + c_1 \leq \theta \leq \theta^*(K) \), capacity will remain fixed but will be fully utilized. Capacity will go unutilized only when \( \theta < \theta(K) \). In Figure 6 this occurs during the intervals \((t_a, t_b)\) and \((t_c, t_d)\).

\(^{16}\)If we allow for depreciation, investment will occur more often and even when demand is below historic highs, but it will still occur in spurts.
The irreversibility of investment induces firms to hold less capacity as a buffer against unanticipated drops in demand. As a result there will be periods of low demands when capacity is fully utilized. A large drop in demand is required for capacity utilization to fall below 100 percent. Most of the time the firm's capacity $K$ will be above $K^*(\theta)$—in Figure 6 exceptions are during the intervals $(t_0, t_1)$ and $(t_2, t_3)$.

The share of the firm's value due to its growth options will also fluctuate with $\theta$. For example, as Table 1 shows, during periods when capacity is growing (so that $K = K^*(\theta)$), this share falls. It also falls when $\theta$ is falling and $K > K^*(\theta)$.

V. The Measurement of Long-Run Marginal Cost

The measurement of long-run marginal cost and its relationship to price are important for industry analyses in general, and antitrust applications in particular. When investment is irreversible, traditional measures understate marginal cost and overstate the amount by which it differs from price, even in a competitive market. This problem is particularly severe when product markets are volatile.

Suppose $\sigma > 0$ and $K = K^*(\theta)$. Using equation (9), the optimality condition $\Delta V(K) = k + \Delta F(K)$ can be written as:

$$\Delta V(K) = b_2 \theta^{p_2} + \theta/\delta - [(2\gamma + c_2)K + c_1]/r$$

$$= k + \Delta F(K)$$

or

$$\theta/\delta - 2\gamma K/r$$

$$= -b_2 \theta^{p_2} + (c_2 K + c_1)/r + k + \Delta F(K).$$
If $\sigma = \alpha = 0$, equation (13') reduces to a more familiar special case:

$$\frac{\theta}{\delta} - 2\gamma K/r = (c_1 + c_2 K)/r + k.$$  

(14)

The left-hand side of (14) is capitalized marginal revenue. (Note that $\theta$ is capitalized at the rate $\delta = \mu - \alpha$ because it is expected to grow at the rate $\alpha$ but is discounted at the risk-adjusted rate $\mu$; if $\sigma = 0$, $\mu = r$.) The right-hand side is full marginal cost: the capitalized operating cost, plus the purchase cost of a unit of capital. Equation (14) is the usual relation between marginal revenue and marginal cost when the former is increasing at a deterministic rate.

Observe that when $\sigma > 0$, two adjustments must be made to obtain full (capitalized) marginal cost, the RHS of (13'). The first term on the RHS of (13') is the value of the firm's option to let the marginal unit of capacity go unutilized, and must be subtracted from capitalized operating cost. The last term is the opportunity cost of exercising the option to invest. In our model the last term dominates the first, so that $K$ must be smaller to satisfy (13'), and marginal cost as conventionally measured will understate true marginal cost.

If the firm is a price-taker, $\gamma = 0$ and $P = \theta$. Price will equal marginal cost, if the latter is defined correctly as in (13'). Unfortunately the first and last terms on the RHS of (13') are difficult to measure, particularly with aggregate data. But if one wishes to compare price with marginal cost, ignoring them can be misleading.\(^{17}\)

VI. Conclusions

Uncertainty over future market conditions affects investment decisions through the options that firms hold—operating options, which determine the value of capital in place, and options to add more capital, which, when investment is irreversible, determine the opportunity cost of investing. By treating capital as homogeneous and focusing on incremental investment decisions, we have tried to clarify the ways in which uncertainty and irreversibility affect the values of these options, and thereby affect the firm's optimal capacity and its market value.

The assumption that firms can continuously and incrementally add capital, though common in economic models, is extreme. Most investments are lumpy, and sometimes quite so. The opposite extreme assumption is that the firm can build only a single plant, and must decide when to build it (the critical $\theta^*$), and how large it should be ($K^*$). As sketched out in the Appendix, this problem can also be solved by the methods used in this paper. For our model: \(i\) the critical $\theta^*$ at which it is optimal to invest increases with $\sigma$; \(ii\) $K^*$ also increases with $\sigma$ (operating options are worth more, and there is only one opportunity to invest); and \(iii\) for every $\theta$, the value of the firm (the value of the single investment option prior to construction, and the value of the plant after construction) is less than it is when the firm can incrementally invest.

Besides ignoring the lumpiness of investment, the model presented here has other simplifying assumptions. It ignores adjustment costs and delivery lags, and includes only one source of uncertainty. It can be extended to account for these factors, but numerical methods may then be required to obtain solutions. Of course once numerical methods are used, other aspects of the model can also be generalized. For example, demand can be a nonlinear function of $\theta$, or $\theta$ could follow some alternative stochastic process, including jumps.

It should be emphasized that the numerical results presented in this paper are based on a specific production technology, and specific functional forms for demand and

\(^{17}\)Robert Hall, 1986, reports that price significantly exceeds marginal cost for most two-digit industries, and finds no explanation for this disparity consistent with competition. Hall's test of marginal cost pricing is based on the relation between the marginal product of labor and the product wage. If firms set marginal operating cost equal to a (constant) proportion of price, his technique will apply, whatever the capital stock. But as shown in Section IV, there can be a wide range of prices for which the firm is capacity constrained, and the ratio of marginal operating cost to price will vary with price.
cost. Also, the assumption that the firm can incrementally invest magnifies the effects of uncertainty, as does the assumption that there is no depreciation. (If capital becomes obsolete rapidly, the opportunity cost of investing will be small.) Thus the quantitative importance of irreversibility and uncertainty may be more limited than the results here would suggest.\footnote{Also, the model examines the investment decisions of a single firm, and ignores entry and competition. If rival firms can appropriate the same investment opportunities, $\delta$ will be larger, which makes the value of the investment option smaller. Steven Lippman and R. Rumelt, 1985, examine the implications of irreversible investment for a competitive market equilibrium, and Avinash Dixit (1987a,b) studies the implications of sunk costs and stochastic prices for firms' entry and exit decisions, and for the reallocation of capital across sectors.}

Subject to these caveats, we find that in markets with volatile and unpredictable demand, firms should hold less capacity than they would if investment were reversible or future demands were known. Also, much of the market value of these firms is due to the possibility (as opposed to the expectation) of increased demands in the future. This value may result from patents and technical knowledge, but it also arises from the managerial expertise, infrastructure, and market position that gives these firms (as opposed to potential entrants) the option to economically expand capacity.

Do firms correctly compute and take into account the opportunity cost of investing when making capacity expansion decisions? Ignoring such costs would lead to overinvestment. John McConnell and Chris Muscarella (1985), have found that for manufacturing firms, market value tends to increase (decrease) when managers announce an increase (decrease) in planned investment expenditures, which is inconsistent with a systematic tendency to overinvest.\footnote{But they find the opposite true for firms in the oil industry, where there may be a tendency to overinvest in exploration and development.} But there is anecdotal evidence that managers often base investment decisions on present values computed with discount rates that far exceed those that would be implied by the CAPM—diversifiable and nondiversifiable risk are sometimes confused, and an arbitrary "risk factor" is often added to the discount rate. It may be, then, that managers use the wrong method to get close to the right answer.

**APPENDIX**

Here we derive equation (9) for $\Delta V(K; \theta)$ and equations (10) and (11) for the optimal investment rule and value of the investment option $F(K; \theta)$.

The value of a marginal unit of capacity, $\Delta V(K; \theta)$, is found by valuing an equivalent "incremental project" that produces 1 unit of output per period at cost $(2\gamma + c_2)K + c_1$, that is sold at price $\theta(t)$, and where the firm can (temporarily and costlessly) shut down if price falls below cost. To value this, create a portfolio that is long the project and short $\Delta V$ units of the output, or equivalently the asset or portfolio of assets perfectly correlated with $\theta$. Because the expected rate of growth of $\theta$ is only $\alpha = \mu - \delta$, the short position requires a payment of $\delta\theta \Delta V$ per unit of time (or no rational investor would hold the corresponding long position). The value of this portfolio is $\Phi = \Delta V - \Delta V \theta$, and its instantaneous return is

$$
(A1) \quad d(\Delta V) - \Delta V \theta dt - \delta \theta \Delta V dt + j[\theta - (2\gamma + c_2)K + c_1] dt.
$$

The last term in (A1) is the cash flow from the "incremental project": $j$ is a switching variable: $j = 1$ if $\theta(t) \geq (2\gamma + c_2)K + c_1$, and 0 otherwise.

By Ito's Lemma, $d(\Delta V) = \Delta V \theta dt + (1/2)\Delta V \theta (d \theta)^2$. Substitute equation (6) for $d \theta$ and observe that the return (A1) is riskless. Setting that return equal to $r \Phi dt = (r \Delta V - r \Delta V \theta) dt$ yields the following equation for $\Delta V$:

$$
(A2) \quad \frac{1}{2} \sigma^2 \theta^2 \Delta V \theta^2 + (r - \delta) \theta \Delta V
$$

$$
+ j[\theta - (2\gamma + c_2)K + c_1] - r \Delta V = 0.
$$

The solution must satisfy the following boundary conditions:

$$
\lim_{\theta \to -\infty} \Delta V(K; \theta) = 0
$$

$$
\lim_{\theta \to \infty} \Delta V(K; \theta) = \theta/\delta - [(2\gamma + c_2)K + c_1]/r
$$

$$
\lim_{\theta \to \infty} \Delta V \theta(K; \theta) = 1/\delta,
$$

and $\Delta V$ and $\Delta V \theta$ continuous at the switch point $\theta = (2\gamma + c_2)K + c_1$. The reader can verify that (9) is the solution to (A2) and its boundary conditions.

Equation (A2) can also be obtained by dynamic programming. Consider the optimal operating policy
(j = 0 or 1) that maximizes the value $\Phi$ of the above portfolio. The Bellman equation is

$$r\Phi = \max_{j = 0,1} \left\{ j\left[ \theta - (2\gamma + c_2) - c_1 \right] - \delta b\Delta V_0 + \frac{1}{dt} E_d d\Phi \right\},$$

that is, the competitive return $r\Phi$ has two components, the cash flow given by the first two terms in the maximand, and the expected rate of capital gain. Expanding $d\Phi = d\Delta V - \Delta V_0 d\theta$ and substituting into (A3) gives (A2).

Finally, note that $\Delta V$ must be the solution to (A2) and the boundary conditions even if the unit of capacity (the "incremental project") did not exist, and could not be included in a hedge portfolio. All that is required is an asset or portfolio of assets (x) that replicates the stochastic dynamics of $\theta$. As Robert Merton (1977) has shown, one can replicate the value function with a portfolio consisting only of the asset x and riskfree bonds, and since the value of this portfolio will have the same dynamics as $\Delta V$, the solution to (A2), $\Delta V$ must be the value function to avoid dominance.

Equation (10) for $\Delta F(K; \theta)$ can be derived in the same way. Using the same arguments as above, it is easily shown that $\Delta F$ must satisfy the equation

$$\frac{1}{2}\sigma^2 \Delta F_{\theta\theta} + (r - \delta) \theta \Delta F_{\theta} - r \Delta F = 0$$

with boundary conditions:

$$\Delta F(K; 0) = 0$$

$$\Delta F(K; \theta^*) = \Delta V(K; \theta^*) - k$$

$$\Delta F_0(K; \theta^*) = \Delta V_0(K; \theta^*),$$

where $\theta^*$ is the exercise point, and $\Delta V(K; \theta^*) - k$ is the net gain from exercising. The reader can verify that equations (10) and (11) are the solution to (A4) and the associated boundary conditions.

The assumption that the firm can invest incrementally is extreme. At the opposite extreme, suppose that the firm can build only a single plant, and must decide when to build it and how large it should be. Now the firm has an option, worth $G(K; \theta)$, to build a plant of (arbitrary) size $K$. Once built, a plant of size $K$ is worth $V(K; \theta) = \int_{-\infty}^{\infty} \Delta V(r; \theta) dr$, where $\Delta V(r; \theta)$ is given by equation (9). It is easy to show that if $K$ and $\theta$ are chosen optimally to maximize $G(K; \theta)$, then $G(K; \theta)$ must satisfy (A4), with $\Delta F$ replaced by $G$. However, the boundary conditions are now:

$$G(K; 0) = 0$$

$$V_K(K^*; \theta^*) - k = 0$$

$$G(K^*; \theta^*) = V(K^*; \theta^*) - kk^*$$

$$G_\theta(K^*; \theta^*) = V_\theta(K^*; \theta^*),$$

where $\theta^*$ is again the exercise point, $K^*$ is the optimal plant size, that is, that $K$ which maximizes $[V(K; \theta^*) - kk^*]$, and $V(K^*; \theta^*) - kk^*$ is the net gain from exercising. Using the first boundary condition, the solution is $G(K; \theta) = a\theta^\beta_1$, where $\beta_1$ is given under equation (9). The constant $a$ and the critical values $\theta^*$ and $K^*$ are determined from the remaining three boundary conditions.

REFERENCES


———, (1987b) "Intersectoral Capital Real-


