MIT Sloan School of Management

MIT Sloan Working Paper 4545-05

June 2005

Sunk Costs and Real Options in Antitrust

Robert S. Pindyck

SUNK COSTS AND REAL OPTIONS IN ANTITRUST[*]

by

Robert S. Pindyck

Massachusetts Institute of Technology
Cambridge, MA 02142

This draft: June 11, 2005

Abstract: Sunk costs play a central role in antitrust economics, but are often misunderstood and mismeasured. I will try to clarify some of the conceptual and empirical issues related to sunk costs, and explain their implications for antitrust analysis. I will be particularly concerned with the role of uncertainty. When market conditions evolve unpredictably (as they almost always do), firms incur an opportunity cost when they invest in new capital, because they give up the option to wait for the arrival of new information about the likely returns from the investment. This option value is a sunk cost, and is just as relevant for antitrust analysis as the direct cost of a machine or a factory.

JEL Classification Numbers: L40; L10, D43

Keywords: Sunk costs, real options, investment decisions, antitrust, entry barriers, market power, mergers.

---

1. Introduction

Sunk costs play a central role in antitrust economics. Sunk costs can be a barrier to entry, for example, which is of obvious importance for merger analysis. In some situations, substantial sunk costs can also make monopoly leveraging, predatory pricing, and other anticompetitive uses of market power feasible, at least in theory. And sunk costs complicate the meaning and measurement of competition in antitrust contexts: If sunk costs are high relative to marginal cost, price will almost surely exceed marginal cost, even though economic profits are zero. In such cases, the locus of competition may be at the entry stage, rather than a later stage of post-entry pricing and production.

It is therefore important that lawyers, judges, and economists involved in antitrust matters understand the meaning of sunk costs, and the issues involved in measuring them. Justifying a merger on the grounds that higher prices would lead to entry, for example, may only make sense if the sunk costs of entry are not too large. And large sunk costs may be a requisite for a claim of monopoly leveraging to make any economic sense. Thus the debates in such cases may revolve around the estimation of sunk costs.

As I will explain in this paper, the measurement of sunk costs is not straightforward. To begin with, many lawyers, business people, and even economists confuse sunk costs with fixed costs, and the distinction between the two can sometimes be quite important. But even after we clarify this distinction, there remain significant measurement issues. First, it is often unclear as to what part of an investment expenditure is in fact sunk. Second, the measurement of sunk costs becomes more complicated when market conditions evolve unpredictably (as they almost always do). In the presence of uncertainty, there is an opportunity cost of investing now, rather than waiting for the arrival of new information about the likely returns from the investment. Firms have options to invest, and when they exercise those options, they give up the associated option value. That option value is also a sunk cost, and is just as relevant for antitrust analysis (and business decision-making) as the direct cost of a machine or a factory. The nature and role of this option value will be the main focus of this paper.

Options to invest (in factories, machines, R&D, marketing and advertising, etc.) are referred to as *real options*, to distinguish them from financial options, such as call and put options on a stock. However, they are closely analogous to financial options. Thus determining the value of

a real option and when it should be optimally exercised can be done using the same techniques that have been developed in the finance literature, and that are widely used on Wall Street.

In the next section I will briefly discuss the distinction between sunk and fixed costs, and the confusion that arises (even in textbooks) over these concepts. I will also discuss measurement issues, e.g., how the possibility of resale affects the "sunkness" of an investment expenditure. In addition, I will explain why the opportunity cost associated with option value is a relevant sunk cost for antitrust analyses.

Section 3 provides an introduction to the theory of option value in the context of capital investments. The basic idea is fairly simple: Firms have *options* to invest in such things as new factories, capacity expansion, or R&D. When a firm invests, it exercises – and thereby gives up – some of these options. Because the firm cannot "unexercise" these options (i.e., it cannot recover its investment expenditures), the value of the options is part of the sunk cost of investing. Of course firms need not exercise their options immediately, and in the presence of uncertainty there is a value to waiting for new information, i.e., a value to keeping these options alive. As mentioned above, these "real options" are analogous to financial options, and can be analyzed accordingly.

In Section 4, I explain in more detail how option value enters in to the sunk cost of investing. I focus on entry decisions, and show how option value can be as much of a barrier to entry as the direct sunk costs that antitrust analyses typically consider. To do this, I use a simple 2-period model, a version of which is first introduced in Section 2. I show that uncertainty over future market conditions effectively "amplifies" the direct sunk cost of entering a market, and can thereby reduce the extent of entry and raise prices.[1] Traditional analyses typically measure the sunk cost of entering a market by adding up the direct costs of, say, building a factory, purchasing equipment, etc. We will see that such an approach underestimates the true sunk cost of entry, and in addition, the magnitude of the true sunk cost depends on the extent of volatility in the particular market.

---

[1] My focus on sunk costs as a barrier to entry in a market with monopoly power is for simplicity, because I am mostly concerned with showing how option value is itself a sunk cost. The same principles apply, for example, to an analysis of how sunk costs can (in theory) make bundling or tying an effective strategy for monopoly leveraging. (See, e.g., Carlton and Waldman (2002).) It is also worth noting that there is disagreement as to what is meant by an "antitrust barrier" to entry, as discussed by Carlton (2004), McAfee et al (2004), and Schmalensee (2004). I follow Schmalensee in using Bain's (1956) definition of an antitrust barrier to entry, i.e., any additional (sunk) cost that an entrant must pay that has already been paid by the incumbent, sufficient to allow the incumbent to raise price without inducing entry.

Section 5 focuses on the way in which information is obtained. The usual assumption is that the firm learns about market conditions simply by waiting. For example, the price of oil fluctuates unpredictably, and to learn what next year's price will be, one can simply wait a year and then look in the newspaper for the current price. This creates an incentive to wait, rather than invest now, because waiting will yield information. However, there are also situations where one learns instead from the actions of others. An example is an R&D program, where a firm might learn about the cost and feasibility of completing the program from the experience of other firms working on similar problems. This also creates an incentive to wait, but now the reason for waiting is to obtain information from the experiences of other firms. This can lead to market failure: the other firms may have the same incentive to wait, with the result that nobody invests. As we will see, this has quite different, but important, implications for antitrust.[2]

I conclude with a summary of the main conceptual points, and further discussion of applications of the theory to antitrust. I discuss, for example, implications of the theory for the analysis of entry barriers, predatory pricing, and merger analysis. I also discuss some of the measurement issues that must be confronted in the application of the theory. The theory of real options has been important in understanding how uncertainty affects investment and industry evolution. My objective is to show that it is also important for understanding and measuring market power and barriers to entry, and evaluating the competitiveness of various practices.

2. Sunk Costs, Fixed Costs, and Option Value.

It is best to begin by clarifying the meaning of a sunk cost, and how it differs from a fixed cost (with which it is sometimes confused).[3] A sunk cost is an expenditure that has been made and cannot be recovered, even if the firm should go out of business. Examples of sunk costs include investments in product development, the construction of a specialized production facility, or an expenditure on advertising. Such expenditures cannot be recovered, and are therefore essentially irrelevant for any ongoing decisions that the firm must make. Of course a

---

[2] In addition, there are situations where a firm can learn only by investing itself. An example is an R&D program undertaken by only a single firm. The cost, and even the feasibility, of completing the program are unknown at the outset, but as the firm invests (i.e., spends money on R&D), the uncertainty is gradually resolved. This creates an incentive to invest early, because investing yields information. I address this case only briefly in the Conclusions; it is analyzed in detail in Pindyck (1993).

[3] For those interested in a textbook discussion of sunk, fixed, and variable costs, it is hard to think of a better reference than Pindyck and Rubinfeld (2005), Chapter 7.

prospective sunk cost is quite relevant for the firm's decisions, which is why sunk costs play an important role in antitrust. A firm might find it uneconomical to enter a market, for example, if entry involves a large prospective sunk cost.

A fixed cost, on the other hand, is an ongoing expenditure. It is independent of the level of output, but it can be eliminated if the firm shuts down. Examples of fixed costs might include the ongoing costs of maintaining the firm's headquarters, the salaries of the firm's top executives, and the costs of auditing the firm's books and preparing financial statements. Fixed costs imply economies of scale (the greater is the firm's output, the lower is its average fixed cost), and can therefore create a barrier to entry.

In general, however, fixed costs tend to be less of a barrier to entry than do sunk costs. Why? Because a fixed cost is a flow of money that does not need to be financed in advance, and that can be terminated if the firm (or plant) shuts down. A sunk cost, on the other hand, is a lump sum payment that must often be made up front before the firm has any significant sales and knows how successful its product will be. Most firms operate in a world of uncertainty, making a lump sum payment riskier than an equivalent (in present value terms) flow of cash that can be terminated should market conditions become unfavorable. The flexibility accorded by a fixed cost thus makes it less of an entry barrier.

Examples of industries with relatively high sunk costs and relatively low fixed costs include computer software, where a large investment must be made in product development before the firm knows how well the product will sell, and copper production, where investments must be made in mines and smelters, and the future price of copper (and thus the return on the investment) is highly uncertain. In both cases, the sunk costs make entry risky and difficult. Airlines are just the opposite. Sunk costs for airlines are low (airplanes can be leased or, if purchased, resold in an active secondary market). In the short run (one to two years) fixed costs are high (the lease payments or, equivalently, the opportunity cost of capital for owning the fleet, cannot be avoided); in the longer run, however, leases can be terminated and planes sold, so most costs are variable. This makes entry relatively easy, as indeed we have seen in recent years.

What about an amortized sunk cost? For example, the cost of a factory might be spread out over ten or fifteen years, perhaps using some accounting measure of depreciation, or some economic measure of depreciation that accounts for obsolescence and deterioration. The annual amortized cost of the factory is often treated as a fixed cost, but this is incorrect. In most cases,

if the firm or factory shuts down, the remaining undepreciated portion of the original cost cannot be recovered. Thus although the amortization of the factory may be useful for accounting purposes, it does not change the fact that the original expenditure for the factory is a sunk cost. Likewise, if the firm, looking forward, is considering whether to build such a factory, it must take into account that the expenditure is a prospective sunk cost.

This distinction between fixed and sunk costs is important, in part because the antitrust implications (e.g., barriers to entry) can be so different. These two types of cost are often confused, particularly when accounting measures of cost are used as proxies for economic costs. For example, an amortized sunk cost (e.g., the depreciation of a factory) is often incorrectly treated as a fixed cost akin to the CEO's salary, and fixed costs are sometimes treated as though they are sunk.[4]

As I discuss below, sunk costs include option value, but fixed costs do not. Firms have discretion over the timing of their sunk costs, and when a firm makes an irrecoverable expenditure, it gives up its option of waiting for more information. Like the expenditure itself, this option value cannot be recovered, i.e., it is sunk. The firm has no discretion, on the other hand, over the timing of its fixed costs. A fixed cost is a flow that must be paid out as long as the firm is operating; it can be terminated only by shutting down.

2.1. What Makes an Expenditure Sunk?

Putting aside option value for the moment, what makes a capital expenditure sunk? A common misconception is that the expenditure is sunk only if the capital purchased cannot be sold to someone else. According to this view, an expenditure of $1 billion on a steel mill would not be sunk, because if market conditions were bad, the mill could be sold to another steel company. But in fact the expenditure *is* sunk, because if market conditions became bad, the value of the mill to any steel company would fall, and no company would be willing to pay the original purchase price.

Investment expenditures are sunk costs when they are *firm- or industry-specific*. For example, most investments in marketing and advertising are firm-specific, and cannot be

---

[4] If there are *no* sunk costs, and if any fixed costs can be immediately eliminated by shutting down, the industry is "contestable" in that there are no entry (or exit) barriers, and "hit-and-run" entry is possible, as pointed out by Baumol, Panzar, and Willig (1982). As Weitzman (1983) has shown, this is equivalent to a horizontal industry supply curve. However, it is difficult to come up with examples of industries that have literally no sunk costs.

recovered. Hence they are clearly sunk costs. A steel plant, on the other hand, is industry-specific -- it can only be used to produce steel. In principle, the plant could be sold to another steel company. If the industry is reasonably competitive, however, the value of the plant will be about the same for all firms in the industry, so there would be little to gain from selling it. For example, if the price of steel falls so that a plant turns out, *ex post*, to have been a bad investment for the firm that built it, it will also be viewed as a bad investment by other steel companies, and the ability to sell the plant will not be worth much. As a result, an investment in a steel plant (or any other industry-specific capital) should be viewed as largely a sunk cost.

Even investments that are not firm- or industry-specific are often partly irreversible because buyers in markets for used machines, unable to evaluate the quality of an item, will offer a price that corresponds to the average quality in the market. Sellers, who know the quality of the item they are selling, will be reluctant to sell an above-average item. This will lower the market average quality, and therefore the market price. This "lemons" problem is common to many markets.[5] For example, office equipment, cars, trucks, and computers are not industry specific, and although they can be sold to companies in other industries, their resale value will be well below their purchase cost, even if they are almost new. Thus expenditures on such equipment should be viewed as at least partly sunk.

2.2. Sunk Costs versus Fixed Costs as Entry Barriers.

I explained above that fixed costs tend to be less of an entry barrier than do sunk costs. A simple two-period example will help to make this clear. This example is an extension of the one used by Schmalensee (2004), with uncertainty introduced.

Suppose there is already one firm in the market (the incumbent monopolist), and a second firm is considering entering. The market demand curve at time t is given by:

$$P_t = \theta_t - Q_t \tag{1}$$

Note from Figure 1 that if θ increases the demand curve shifts out to the right; if it decreases the demand curve shifts to the left. In the first period ($t = 1$), firms know that $\theta_1 = 10$, but they do not know what its value will be in period 2; they only know that it will equal 0 or 20, each with

---

[5] This implication of asymmetric information was first analyzed by Akerlof (1970). For a textbook discussion, see Pindyck and Rubinfeld (2005).

probability ½. (Firms learn the value of θ_2 in period 2.) I will assume that both the incumbent

monopolist and the potential entrant have zero marginal cost. I will also assume that should

entry occur so that there are two firms in the market, competition between the firms would lead

to a Cournot equilibrium.[6] It can be easily shown that in the Cournot equilibrium each firm

produces a quantity $q_1 = q_2 = \theta/3$, so that the total quantity produced is $2\theta/3$, the market price is

$\theta/3$, and each firm earns a profit of $\theta^2/9$. Thus in period 1 each firm earns a profit of 100/9, and

in period 2 each firm earns either 0 or 400/9, depending on the outcome for θ_2.

Suppose that entry can only occur in the first period, and it requires the payment of a sunk

cost S. How large would that sunk cost have to be to deter entry? Assuming a discount rate of

zero, the entrant's expected total net value of profits in the two periods is:

$$\text{NPV} = \frac{100}{9} + \frac{1}{2}(0) + \frac{1}{2}\left(\frac{400}{9}\right) - S = \frac{300}{9} - S \qquad (2)$$

Thus entry would occur as long as the sunk cost is less than 300/9. Note that if this maximum

sunk cost were amortized over the two periods, it would be equivalent to a commitment to pay

300/18 in each period.

Now suppose that instead of a sunk cost, entry involves the payment of an annual fixed cost

F. How large would this fixed cost have to be to deter entry? In period 2, if θ turns out to be 0,

the firm will shut down and avoid the fixed cost. As a result, the expected total net value of the

profits in the two periods is now

$$\text{NPV} = \left(\frac{100}{9} - F\right) + \frac{1}{2}(0) + \frac{1}{2}\left(\frac{400}{9} - F\right) = \frac{300}{9} - \frac{3}{2}F \qquad (3)$$

Thus entry would occur as long as F was less than 200/9. This is larger than the maximum

amortized sunk cost of 300/18 paid in each of the two periods (regardless of θ_2.) Entry can

occur with a larger fixed cost, because if market conditions in period 2 are unfavorable (i.e., $\theta_2 =$

0), the firm can shut down and avoid the fixed cost in that period. This flexibility makes a fixed

cost less of an entry barrier than a sunk cost.

---

[6] See Chapter 12 of Pindyck and Rubinfeld (2005) for a textbook exposition of the Cournot model.

2.3. Option Value as a Sunk Cost.

Now suppose that entry can occur in either the first or the second period, and that, as before, it requires the payment of a sunk cost S. We will set S at 300/9, which, as we saw from eqn. (2), is the value that makes the NPV of entry in period 1 just equal to zero. But what if instead the potential entrant waits until period 2 before deciding whether to enter? In this case, he will enter only if $\theta_2 = 20$, so that his profit in period 2 is 400/9. The probability that θ_2 will equal 20 is 1/2, so the NPV (as of period 1) in this case is $(1/2)(400/9 - S) = (1/2)(400/9 - 300/9) = 100/18$. This NPV is higher (100/18 versus 0), so clearly it is better for the entrant to wait until period 2 before deciding whether to enter.

Wouldn't the entrant also be justified in entering in period 1, given that the NPV of that strategy is just zero? No. The NPV as calculated in eqn. (2) includes only direct costs, ignoring an important opportunity cost. The value of the firm's option to enter the market is the NPV that it can obtain by using an optimal strategy, i.e., it is 100/18. If the firm enters in period 1, it gives up this option value (the value of waiting for information about θ_2). This lost option value cannot be recovered, and is thus a sunk cost. The true sunk cost of entering in period 1 is therefore $300/9 + 100/18 = 700/18$, which means that the true NPV of entering in period 1 is negative. By excluding the lost option value from the estimate of its sunk cost, the firm would underestimate that cost, and incorrectly think that it is economical to enter in period 1. Likewise, an antitrust analysis concerned with the possibility of early entry by another firm would underestimate the sunk cost "barrier" to such entry.

In the preceding example, I distinguished between an incumbent firm and a potential entrant. It is important to keep in mind, however, that in an uncertain environment the distinction between sunk and fixed costs and the existence of option value have broader implications for industry equilibrium. In general, the higher are the sunk costs required for entry (whether those sunk costs are direct or are opportunity costs associated with option value), the smaller is the number of firms we can expect in equilibrium.

The example discussed above is simplified and somewhat artificial. In the next section, I provide a brief introduction to the theory of real options, and explain in more detail why option value is a sunk cost.

3. Real Options.

Most analyses of capital investment, in antitrust applications and elsewhere, are based on a simple investment rule that has been taught widely in business schools, and is the foundation for much of neoclassical investment theory in economics: the Net Present Value (NPV) rule. This rule says that a firm should invest in a project if the NPV of the project is positive, i.e., if

$$\text{NPV} = -I_0 - \frac{I_1}{(1+\rho)} - \dots + \frac{\pi_1}{(1+\rho)} + \frac{\pi_2}{(1+\rho)^2} + \dots > 0 \qquad (4)$$

where I_0, I_1, … are investment outlays, π_1, … are net cash flows arising from the investment, and ρ is the discount rate, usually the firm's weighted average cost of capital (WACC). If the investment is completely reversible (i.e., the investment could be "undone" and the expenditure recovered), or if there is no uncertainty over the future cash flows, or if this investment is a now-or-never proposition (i.e., there is no possibility of delaying the investment), then this rule is correct. However, if the investment is fully or partly irreversible (sunk), there is uncertainty over the cash flows, and the investment could be delayed, the rule is wrong. In particular, the use of this rule does not maximize the firm's value, i.e., the firm would do better using a different rule.

Why is this NPV rule incorrect? Because it makes the wrong comparison – it compares investing today with never investing. The correct comparison is investing today versus waiting, and perhaps (depending on how market conditions turn out) investing at some unspecified time in the future. Put differently, a firm with an opportunity to invest is holding an "option" analogous to a financial call option – it has the right but not the obligation to buy an asset at some future time of its choosing. When a firm makes an irreversible investment expenditure, it exercises its option to invest. It gives up the possibility of waiting for new information to arrive that might affect the desirability or timing of the expenditure; it cannot disinvest should market conditions change adversely. This lost option value is an opportunity cost that must be included as part of the total cost of the investment. As a result, the NPV rule "Invest when the value of a unit of capital is at least as large as its purchase and installation cost" must be modified. The value of the unit must exceed the purchase and installation cost, by an amount equal to the value of keeping the investment option alive.[7]

---

[7] One of the first studies to make this point was McDonald and Siegel (1986). For a detailed treatment of the value of waiting, and real options in general, see Dixit and Pindyck (1994.) For a shorter overview, see Pindyck (1991).

3.1. A Simple Example.

A simple example may help to clarify these ideas.[8]   Suppose a firm is considering an investment in a new factory that will cost $10,000, and will immediately generate annual cash flows that will continue forever.   This year, the cash flow will be $1,000, but next year, depending on market conditions, it will either increase to $1,500 or decrease to $500, with equal probability. For simplicity, let us assume that the cash flow will then stay at that level ($1,500 or $500) for all future years.

What is the NPV of this investment, assuming the firm invests immediately?   Since the expected value of the cash flow from next year on is $1,000, the NPV is:

$$NPV = -10,000 + \sum_{t=0}^{\infty} \frac{1000}{(1+\rho)^t} \tag{5}$$

where ρ_t is the discount rate, i.e., the cost of capital. Assuming the value of this discount rate is 10 percent, the NPV is equal to $-10,000 + 11,000 = \$1000$. The NPV is positive, so investment seems warranted. But is it?

Now suppose the firm waits a year, and then invests only if the annual cash flow goes up to $1500. Since the probability of this happening is 0.5, the NPV as of today is:

$$NPV = (0.5) \left[\frac{-10,000}{1.1} + \sum_{t=1}^{\infty} \frac{1500}{(1.1)^t} \right] = \frac{3250}{1.1} = \$2955 \tag{6}$$

By waiting a year before deciding whether to invest, the NPV is $2955, whereas it is only $1000 if the firm invests today. Clearly it is better to wait rather than invest now – even though the NPV of investing now is positive. The reason is simple: By waiting, the firm can avoid the consequences of an unfavorable outcome. Had the firm invested today and had the cash flow dropped to $500 per year, the *ex post* NPV would be

$$NPV = -10,000 + 1000 + \sum_{t=1}^{\infty} \frac{500}{(1.1)^t} = -10,000 + 6000 = -\$4000 \tag{7}$$

In other words, the firm would find itself losing money. If the firm waits, on the other hand, it would simply not invest if the cash flow fell to $500.

What is the value of having the ability to wait, rather than facing a now-or-never decision? It is just the difference in the two NPVs that we calculated above, i.e., $2955 – $1000 = $1955.

---

[8] This is a modified version of an example in Chapter 2 of Avinash K. Dixit and Robert S. Pindyck, *Investment Under Uncertainty*, (Princeton: Princeton University Press; 1994).

This is the value of the "flexibility option." In other words, the firm should be willing to pay up to $1955 more for an investment opportunity that is flexible than one that only allows the firm to invest now.

We have seen that the NPV of investing today is positive, so what are we missing? The problem is that the NPV as calculated in eqn. (5) does not properly account for the full cost of investing today – it only accounts for the direct expenditure of $10,000. It ignores the *opportunity cost* of giving up the firm's option to wait for more information. That opportunity cost is just equal to the value of the option when it is optimally exercised, i.e., when the firm waits rather than invests immediately. That value is the NPV today when it waits, i.e., it is $2955. Thus the true NPV of investing today is $1000 – $2955 = –$1955. Hence the NPV of investing today, when properly calculated so as to include opportunity costs, is *negative*.

The key point here is that an (optimal) investment decision must take into account the full sunk cost of investing – the direct cost ($10,000 in this example) plus the opportunity cost of giving up the firm's option to invest (which in period 1, before the change in the value of the factory is know, is $2955). In this example, the payoff from the investment, i.e., the value of the completed factory is $11,000 in period 1, which exceeds the direct sunk cost, but not the full sunk cost. In period 2, of course, there is no further uncertainty, and thus there is no longer any option value, so the decision to invest can be based solely on a comparison of the $10,000 direct sunk cost with the payoff from investing ($16,500 if the annual cash flow increases to $1500, but only $5,500 if the annual cash flow decreases to $500).

This simple example, of course, is in some ways quite unrealistic. Perhaps most importantly, we assumed that all of the uncertainty gets resolved in one year, i.e., next year the annual cash flow will either increase or decrease, but then it will no longer change from this high or low value. In reality, there is always uncertainty over future cash flows. Market conditions are constantly evolving, so that the cash flows from a capital investment will likewise constantly evolve. In such a situation, one must solve an option pricing problem to determine the value of the firm's option to invest and its optimal investment decision. Fortunately, methods developed in finance can be brought to bear, so that solving this problem is usually quite feasible.

3.2. The Option to Invest.

In the business world, future market conditions are almost always uncertain. A firm must

take into account that demand for its product, cost conditions, and thus the value of the factory that produces the product will continually fluctuate, and those fluctuations are in large part unpredictable. Some uncertainty will be resolved a year from now – demand might turn out to be stronger or weaker than anticipated, for example – but there will still be uncertainty looking down the road another year, or three or five years.

The nature of this ongoing, evolving uncertainty is perhaps easiest to understand in the context of oil prices, and an investment in an oil producing facility. Suppose an oil company owns an undeveloped offshore oil reserve. (This means that it knows that there is a certain amount of oil under the ground, but it cannot currently extract and sell the oil.) To develop the reserves will cost $1 billion. Development will result in the ability to produce a steady stream of oil over, say, the next ten years. Should the company spend the $1 billion and develop the reserve now, or wait to see if the price of oil increases or decreases? Suppose in addition that a conventional NPV analysis, based on an expected future price of oil that increases at 2 percent per year from today's price, gives an NPV of $10 million.

Although the NPV is positive, given the volatility of oil prices it is unlikely that development today is economical (in the sense of value-maximizing). The key here is that the firm has an option to invest in development. It has what is essentially a call option on a dividend paying stock. Table 1 shows the analogy between the financial call option, and the "real option" to develop the oil reserve. The value of a financial call option depends on five key variables: the price of the stock on which the option is written, the exercise price, the time to expiration, the volatility of the stock price, and the dividend rate for the stock. Note that the value of the call option does *not* depend on a forecast of the future price of the stock. In the same way, the value of an undeveloped reserve depends in on the value of the developed reserve, the cost of development, the time to expiration (which might be infinite, i.e., the option never expires, or, as with offshore leases in the U.S., might correspond to a relinquishment requirement of several years), the volatility of the value of reserve once it has been developed (which would usually correspond to the volatility of the price of oil), and the net payout or profit rate (which for a developed oil reserve is net production revenue less depletion). As with the financial call option, the value of the undeveloped reserve does *not* depend on a forecast of the future value of a developed reserve (or the future price of oil).

Table 1: Comparison of a Call Option with Two Investment Opportunities

Call Option on Stock	Undeveloped Oil Reserve	Option to Build Factory
Price of stock	Value of developed reserve	Value of completed factory
Exercise price	Cost of development	Cost of building factory
Time to expiration	Relinquishment requirement (if there is one)	None (perpetual option)
Volatility of stock price	Volatility of value of developed reserve	Volatility of value of completed factory
Dividend on stock	Net production revenue from developed reserve less depletion	Net cash flow from completed factory

An undeveloped oil reserve is thus exactly analogous to a call option on a dividend-paying stock. The only significant difference is that financial call options are typically short-lived (several months), whereas the option to develop an undeveloped reserve is typically long-lived (several years, or even perpetual). The long life of the real option makes it easier to evaluate, because the time element can be largely ignored.

When valuing a financial call option, one needs some description (or model) of the behavior of the price of the stock on which the option is written. That behavior is often best described as a *geometric random walk*: At each small interval of time (e.g., each minute) there occurs a random increase or decrease in the price, such that the *percentage change* in the price is normally distributed, and such that the random change at each instant is independent of the change in the previous instant. This means that random changes in the price of a stock are unpredictable, so that "technical analyses" of the behavior of the price in the past are of no value in predicting what the price will do in the future. Although this description does not fit the actual behavior of stock prices perfectly, it is an excellent first approximation, and is usually the basis for valuing stock options and other derivatives by financial institutions.[9]

---

[9] It is the basis, for example, of the Black-Scholes formula for the value of a call option. To be more precise, the price of the stock is assumed to follow a *geometric Brownian motion*, so that no matter how short the time intervals, random percentage changes in the price are independent of each other, and so that the variance of the future price grows linearly with the time horizon. See Chapter 3 of Dixit and Pindyck (1994), Hull (1997), or McDonald (2004) for a much more detailed explanation.

The value of a developed oil reserve will depend directly on the price of oil, which likewise can be described – at least to a first approximation – as a geometric random walk. Daily changes in the price of oil are generally independent of changes the previous day (or previous week, month, etc.), so that future prices cannot be predicted based on past behavior. Thus the same methods used for financial options can be applied to the valuation of an undeveloped reserve, and the determination of the optimal point at which the reserve should be developed.[10]

Table 1 also shows the very similar analogy between a financial call option and the real option to build a hypothetical widget factory. (I am assuming here that the factory would have some fixed production capacity, and I am ignoring such complications as the amount of time it would take to complete the factory once construction has started.) In this case it is the value of a completed factory that corresponds to the price of the stock, and the cost of construction that corresponds to the exercise price. Presumably, the completed factory would be expected to yield a positive net cash flow (i.e., revenue less variable cost), which corresponds to the dividend on a stock. But assuming that the factory could be built at any time in the future, the option to invest is perpetual, unlike a financial option.

The value of the widget factory will likely depend on a number of variables that evolve over time unpredictably, e.g., wage rates and costs of materials. But the variable that is likely to be the most important determinant of the value of the factory is the price of widgets, which will also evolve unpredictably. Because the volatility of the value of the factory will depend mostly on the volatility of the price of widgets, it is important to describe the way in which this price evolves over time. Depending on how competitive is this hypothetical widget market, the price might be best described as a random walk, or alternatively as a mean-reverting process, in which the price can evolved randomly but tends to revert to a mean or "normal" level, which would typically be long-run marginal cost. For simplicity, I will assume that the price of widgets, like the price of oil, follows a geometric random walk.[11] This would be realistic if firms in the

---

[10] See Paddock, Siegel, and Smith (1988) and Chapter 12 of Dixit and Pindyck (1994) for detailed discussions of the undeveloped oil reserve problem.

[11] If the price follows a geometric random walk, the percentage change in the price each period (e.g., each week or month) is a normally distributed random variable. An important advantage of the geometric random walk assumption is that it greatly simplifies the problem of solving for the value of the option to invest and finding the optimal timing rule for investment. However, solutions are also feasible for mean-reverting price processes, and these are widely used in real option problems. See Dixit and Pindyck (1994) for a detailed discussion and several examples.

market have considerable market power (which would presumably be the case in antitrust applications), so that price can deviate considerably from marginal cost, and is driven more by fluctuations in market demand.

As we saw in Table 1, the option to invest in this widget factory will depend on the level and the volatility of the value of the completed factory, the cost of construction, the expected cash flow from the factory, and (not shown in the table) the risk-free interest rate. Assuming that the factory can be built quickly and that the price of widgets and the value of the completed factory follow a geometric random walk, the value of the option to invest and the optimal investment rule are easy to calculate, and are shown graphically in Figure 2.[12] In that figure, the value of the option to invest, $F(V)$, is plotted against the value of the completed factory, V, assuming that the cost of building the factory (i.e., the exercise price of the option) is $I = \$100$ million. The value of the option depends on the volatility (denoted by σ) of the value of the factory; I have plotted the option value for two different levels of volatility, σ_1 and $\sigma_2 > \sigma_1$. Note that a higher level of volatility implies a higher option value.[13]

In addition to the value of the option, the graph also shows the line $V - 100$, which is the net payoff from building the factory as a function of V. The conventional NPV rule would say to build the factory as long as this net payoff is positive, i.e., as long as $V > 100$. But note that at the point where $V = 100$, the value of the option is much larger than the net payoff of zero – it is about \$25 million for the lower level of volatility, and about \$35 million for the higher level. Suppose that V were \$101 million. Following the conventional NPV rule and building the factory (i.e., exercising the option) would yield a net payoff of \$1 million. However, the firm would be giving up its option to invest, which is worth about \$25 or \$35 million, depending on the level of volatility. It is clearly not optimal to build the factory at this point.

At what value V would it indeed be optimal to go ahead and build the factory? Consider the lower level of volatility, σ_1. Note from the graph that the value of the option, $F(V, \sigma_1)$, just

---

[12] The fact that the price of widgets follows a geometric random walk does not necessarily imply that the value of the widget factory also follows a geometric random walk. If, for example, marginal cost is substantial and the factory can be temporarily shut down or output temporarily reduced when the price falls below marginal cost, the value of the factory will follow a more complicated process. In such a case the firm that owns the factory would have "operating options," i.e., options at each point in the future to temporarily reduce or terminate production. Such complications are discussed in detail in Dixit and Pindyck (1994), but I ignore them here for simplicity.

[13] In calculating $F(V)$, I assumed that the risk-free rate of interest is 4 percent, and the annual net cash flow from the factory is 4 percent of the factory's value. The volatilities, expressed as the standard deviation of annual percentage changes in V, are $\sigma_1 = 0.2$ and $\sigma_2 = 0.3$.

becomes equal to the net payoff $V - 100$ at the point $V_1^* = 200$. At this point, the net payoff from building the factory, $200 - 100 = 100$, is just large enough to compensate the firm for giving up (by exercising) its option to invest. Put differently, at $V = 200$, the *full* cost of the factory, i.e., the direct cost (100) plus the lost option value (100) is just equal to the payoff (200). For values of V above 200, it is always optimal to exercise the investment option, so the value of the option $F(V)$ and the net payoff $V - 100$ are identical. Note that a modified NPV rule, which included the lost option value as part of the cost, would give correct guidance for the investment decision.

What, then, is the sunk cost of building this widget factory? Economic costs include any relevant opportunity costs. In this case there is a sunk opportunity cost, which is the loss of option value that occurs when the firm goes ahead and invests. Thus the sunk cost of building the factory is the direct cost of construction plus the lost option value, i.e., \$200 million. If one were concerned with sunk costs as a barrier to entry, this would be the relevant number to consider in the context of this market.

Now suppose that the widget market is more volatile than we had assumed, i.e., the volatility is $\sigma_2 > \sigma_1$. As Figure 2 shows, the value of the option to invest is now higher. In addition, the critical value V_2^* at which point it is economical to invest is now higher: about \$272 million instead of \$200 million. The reason is that with more volatility, there is a greater opportunity cost of investing now rather than waiting for more information. But this means that *the full sunk cost of building the factory is now higher* (it is now \$272 million), and *much* higher than the \$100 million direct cost of construction. Now the opportunity cost component of the full sunk cost is greater than the direct construction cost component.

The antitrust implications of this result may be counterintuitive at first, but they are profound: Sunk cost barriers to entry can depend, often to a considerable extent, on the degree of market volatility and uncertainty. Analyses based only on direct sunk costs may greatly underestimate the true sunk cost barriers to entry. Furthermore, the extent of volatility is a basic structural feature of a market that plays an important role in determining the number of firms that can be expected to enter and compete.

So far we have only considered the investment decision of a single firm. Our emphasis is on entry, and the extent to which higher prices charged by a monopolist will induce new firms to enter the market, thereby bringing prices back down. I turn to that next.

4. Sunk Costs as a Barrier to Entry.

To see how option value can be a barrier to entry, let's return to the simple two-period example of entry that was introduced in Section 2. Recall that the market demand curve in that example was $P_t = \theta_t - Q_t$, where θ_t was assumed to equal 10 in period 1, would become either 0 or 20 in period 2, and then would remain at that value for all future time. To make this a bit more general, I will instead assume that θ_2 will equal either $10 + \varepsilon$ or $10 - \varepsilon$, each with probability ½. Thus ε is a measure of uncertainty. (In the previous example, we fixed ε at 10.) Note that the variance of θ_2 is ε^2, so by changing ε we can change the amount of uncertainty over demand in period 2, and see how this affects the entry decision.

I will assume that a monopolist is already in the market. Any number of additional firms can enter (in period 1 or period 2), but to do so, each must pay a direct sunk cost S. Let n be the number of firms that enter (so that post-entry there are $n + 1$ firms in the market). Because of uncertainty over market conditions in period 2, the full sunk cost of entering will exceed the direct cost S. One of our objectives is to determine how n, and the resulting market price P, depend on S and ε.

I will further assume that entry takes no time, that any entrant can remain in the market forever with no further expenditure (i.e., there is no depreciation), and that marginal cost is zero. Thus if a firm enters in period 1, its NPV of entry is the profit it earns in that first period, plus the sum of its expected discounted profits in all future periods, minus the direct sunk cost S. I will assume that all firms have a discount rate of 10 percent. Finally, I will assume that post entry, competition among firms is Cournot.

Suppose that n firms have entered the market, so that including the original monopolist, there are a total of $n + 1$ firms competing. Then it is easy to show that the Cournot equilibrium is that each firm produces a quantity $Q_i = \theta/(n+2)$ in each period, so that the total quantity produced is $Q = \theta(n+1)/(n+2)$, and the market price is $P = \theta/(n+2)$. Also, each firm earns a profit of:

$$\pi_i = \theta^2/(n+2)^2 \tag{8}$$

and total consumer surplus in each period is:

$$\mathrm{CS} = \frac{1}{2}(\theta - P)Q = \frac{1}{2}\left(\frac{n+1}{n+2}\right)^2 \theta^2 \tag{9}$$

To show how option value contributes to the full sunk cost of entry, we will begin by assuming that entry can only occur in the first period, and determine the number of firms that will enter, which we denote by n_1. We then take this number of firms has fixed, and ask whether they would prefer to wait until period 2 before deciding whether to enter.[14] We will see that they would indeed prefer to wait, and should they wait, they would only enter if θ increases, i.e., if $\theta_2 = 10 + \varepsilon$. The option value associated with entry is what each firm would give up (in the form of a reduction in the NPV of entry) by entering in period 1 instead of waiting. To determine the contribution to sunk cost, we find the sunk cost $S_2 > S$ that makes the NPV when the firms wait equal to the NPV when the firms (facing the original sunk cost S) enter in period 1. We will examine the "markup" over the direct sunk cost, measured by the ratio S_2/S, and see how it depends on ε, i.e., on the degree of uncertainty.

Entry in Period 1. Suppose all entry must occur in the first period. In this case, how many firms will enter? The NPV for each firm that enters is:

$$
\begin{aligned}
\text{NPV}_i^1 &= \pi_{i1} + E\sum_{t=2}^{\infty} \frac{\pi_{it}}{(1+r)^{t-1}} - S \\
&= \frac{100}{(n+2)^2} + \frac{100+\varepsilon^2}{(n+2)^2}\sum_{t=2}^{\infty}\frac{1}{(1+r)^{t-1}} - S \\
&= \frac{1100+10\varepsilon^2}{(n+2)^2} - S
\end{aligned}
\tag{10}
$$

Here, NPV_i^1 denotes the NPV for firm i when entry occurs in period 1. Any firm can enter, so entry will continue up to the point that the NPV for each firm is zero. Setting NPV_i^1 in eqn. (10) equal to zero and solving for n gives us the number of firms that will enter in period 1:

$$
n_1(S,\varepsilon) = \sqrt{\frac{1100+10\varepsilon^2}{S}} - 2
\tag{11}
$$

As we would expect, n_1 is lower the larger is the sunk cost of entry, S. But in addition, n_1 is *higher* the larger is ε, i.e., the greater is the extent of uncertainty over θ_2. The reason is that the profit for each firm is a convex function of θ_2; note from eqn. (8) that π_i depends on the square of

---

[14] Of course if firms wait until period 2 before deciding whether to enter, the number that actually do enter will differ from n_1. We keep the number of firms fixed at n_1 in order to isolate the value of waiting, which is the option value given up when firms enter in period 1.

θ_2, so the increase in profit from an increase in θ_2 is greater than the decrease in profit from a decrease in θ_2 of the same magnitude.

The Value of Waiting. Now, take this number of firms, $n_1(S, \varepsilon)$, as fixed, and consider whether the firms would prefer to wait until period 2 before making the decision to enter. Recall that we found n_1 by setting the NPV in eqn. (10) equal to zero. Thus the firms would prefer to wait if the expected NPV when they wait, calculated as of period 1 (before they observe the outcome for θ_2), is greater than zero. It is easy to show that if they wait and $\theta_2 = 10 - \varepsilon$, the NPV of entering at that point would be negative (for any positive value of ε), so entry would not occur. Thus entry would only occur if $\theta_2 = 10 + \varepsilon$, in which case the NPV of entry (discounting back to $t = 1$) would be:

$$\mathrm{NPV}_i^2(\theta_2 = 10 + \varepsilon) = \frac{11(10 + \varepsilon)^2}{(1.1)(n_1 + 2)^2} - S \tag{12}$$

The probability that $\theta_2 = 10 + \varepsilon$ is one-half, so the expected NPV at $t = 1$ (before θ_2 is known) is:

$$\mathrm{NPV}_i^2 = \frac{5(10 + \varepsilon)^2}{(n_1 + 2)^2} - .5S \tag{13}$$

Now substitute eqn. (11) for n_1:

$$\mathrm{NPV}_i^2 = \frac{1}{2} S \left[\frac{(10 + \varepsilon)^2}{110 + \varepsilon^2} - 1 \right] \tag{14}$$

This NPV is greater than zero as long as ε is greater than 0.5. (If ε is less than 0.5, the loss from discounting at 10% over one period exceeds the expected gain from waiting to obtain information about θ_2.) Thus, assuming that $\varepsilon > 0.5$, these firms would prefer to wait until period 2 before making their entry decision.

The Full Sunk Cost of Entry. We can now determine the option value associated with the sunk cost of entry. To do so, we ask what the firms give up by entering at $t = 1$ rather than waiting until $t = 2$. Equivalently, we find the sunk cost $S_2 > S$ that makes the NPV of waiting until $t = 2$ just equal to the zero NPV when n_1 firms, each facing a sunk cost S, enter at $t = 1$. That is, we find the cost S_2 that makes $\mathrm{NPV}_i^2(n_1, S_2) = \mathrm{NPV}_i^1(n_1, S) = 0$. Using eqn. (13), S_2 is given by:

$$\frac{5(10 + \varepsilon)^2}{(n_1 + 2)^2} - .5S_2 = 0 \tag{15}$$

Once again, substituting eqn. (11) for n_1:

$$\frac{1}{2}S\frac{(10+\varepsilon)^2}{110+\varepsilon^2}-\frac{1}{2}S_2=0 \qquad (16)$$

Finally, rearrange eqn. (16) to find the ratio S_2/S:

$$S_2/S=\frac{(10+\varepsilon)^2}{110+\varepsilon^2} \qquad (17)$$

Note that $S_2/S>1$ as long as $\varepsilon>0.5$.

Eqn. (17) translates the option value that is lost when each of the firms enters the market into an equivalent "markup" over the direct sunk cost of entry. S_2 is the full sunk cost of entry, i.e., the direct sunk cost S plus the option value that is lost by irreversibly investing. As illustrated in Figure 3, the greater is the degree of uncertainty, the greater is this lost option value, so that S_2/S increases as ε increases.[15] One way to understand this is in relation to Figure 2, which was discussed in the previous section. In that figure, the direct sunk cost of investing was $100, but (for volatility σ_1) the option value is $100, so that the full sunk cost is $200. At a higher level of volatility (σ_1), the option value is larger, and so is the full sunk cost.

As can be seen from this example (and from the discussion in Section 3), the full sunk cost that is relevant to an entry decision (and therefore relevant to an analysis of entry barriers) is greater than the direct sunk cost that is typically considered. Furthermore, the magnitude of the full sunk cost depends on the volatility of market conditions. Measuring sunk costs in a meaningful way must therefore include an analysis of volatility in the particular market.

5. Learning from Others.

In all of the examples we have examined so far, firms learn about market conditions simply by waiting, i.e., they learn from "nature." There are situations, however, where one learns instead from the actions of others. An example is an R&D program, where a firm might learn about the cost and feasibility of completing the program from the outcomes of other firms working on similar problems. Another example is the exploration for oil and gas reserves. Suppose four firms own leases that give them the rights to explore (and, if the exploration is successful, ultimately produce oil or gas) on various different offshore tracts. Because outcomes

---

[15] If $\varepsilon<0.5$, the value of waiting is negative, i.e., the firm's NPV is larger if it invests immediately.

across tracts are correlated, each firm could obtain information about the likely outcome of exploration from the outcomes of the other firms.

In situations like these, there is also an incentive to wait, but now the reason for waiting is to obtain information from the experiences of other firms. This can lead to market failure: the other firms may have the same incentive to wait, with the result that nobody invests. As we will see, this has quite different, but important, implications for antitrust.

Let's return to the example developed in Section 3.1, in which a firm considers an investment in a new factory that will cost $10,000, and will immediately generate annual cash flows that will continue forever. Recall that in the example, the cash flow this year will be $1,000, but next year it will either increase to $1,500 or decrease to $500, with equal probability, and will then stay at that level ($1,500 or $500) for all future years. We will change the example so that *two firms* are considering investing in a factory. We will assume that the market is large enough so that the cash flow to each firm is independent of the sales of the other firm. One other change we will make is to assume that the cash flow from a factory is either $1,500 or $500 in *every year*, again with equal probability. The only way a firm can learn whether the cash flow will be $1,500 or $500 is by investing itself, *or* by observing the outcome when the other firm invests. Now, let's examine the incentives for each of the firms.

For both firms, the NPV of investing now (again using a 10 percent discount rate) is:

$$\text{NPV}_i^{\text{NOW}} = -10,000 + \sum_{t=0}^{\infty} 1000/(1.1)^t = -10,000 + 11,000 = \$1000 \qquad (18)$$

Suppose that Firm 2 is going to invest now. Should Firm 1 wait a year before deciding whether to invest? If it does wait, it will only invest if it learns (from the experience of Firm 2) that the cash flow is $1500. Hence, assuming that Firm 2 will indeed invest now, the NPV for Firm 1 if it waits is:

$$\text{NPV}_1^{\text{WAIT}} = \frac{1}{2}\left[\frac{10,000}{1.1} + \sum_{t=1}^{\infty} 1500/(1.1)^t\right] = \$2955 \qquad (19)$$

In this situation, it is clearly better to wait.

The problem is that Firm 2 is thinking the same thing, and would like to wait for Firm 1. Suppose that, as a result, neither firm invests now, and both wait a year, hoping (in vain) that the other firm will invest. If at the end of the year both firms then go ahead and invest (without the

benefit of any knowledge about the cash flow), the NPV for each firm, calculated as of today, will be:

$$NPV_i^{WAIT} = 1000/1.1 = \$909 \tag{20}$$

We thus have a gaming situation, the payoffs for which are shown below in the form of a payoff matrix. Note that each firm would like the other one to invest first, but in all likelihood both firms will wait a year, and in the end gain nothing in the way of new information. If there were no further possibility of waiting, they would then both invest, and be worse off than they would have had they simply invested now. (Of course, consumers would likewise be worse off.) Also note that if the firms could collude, they would probably agree to toss a coin to see who will invest first. In that case, each firm would have an expected NPV of ½(1000) + ½(2955) = $1977.50.

Firm 2

	Now	Wait
Now	1000, 1000	1000, 2955
Wait	2955, 1000	909, 909

Firm 1 (label at left, rows Now/Wait)

Assuming that collusion and agreement to toss a coin is not possible, how should we expect the two firms to play this game? These firms are effectively in a *war of attrition*; each firm is hoping the other will "blink" first. As long as each firm thinks that the other is reasonably likely to move first, neither firm will invest. And of course there is no reason for this process to stop at the end of one year. Suppose that both firms decide to wait a year. At the end of the year, both observe that no one has invested. They are now in the very same situation, and it is quite possible that once again, neither firm will invest.

If there are four or five firms in this situation, rather than only two, it is even more likely that no one will invest. Why? Because if any one firm invests, all of the others will benefit. Each firm now figures that it is likely that at least one of the several other firms will go ahead and invest, and thus will rationally conclude that it is better to wait. Thus it is possible for a long

time to pass with none of the firms investing – all of them holding out for the possibility that at least one of the others will invest first. The firms all lose in this situation, as do consumers.

What are the antitrust lessons from this? The basic problem here is that there is an externality: When a firm invests, other firms in the industry benefit from this investment, but the firm doing the investing does not capture these benefits. This externality can lead to market failure in the form of inefficient underinvestment. Clearly if there were some way to coordinate the investment decisions of the various firms, investment would occur sooner, to the benefit of consumers. In the case of oil and gas exploration, the coordination often takes the form of unitization, i.e., pooling the leases (or other resource holdings) of the various firms, and then exploring and developing the reserves as a joint venture. Given the benefits involved, it is hard to see how this should create problems from the point of view of antitrust. In the case of R&D, when learning across firms is important, there are clear incentives for joint ventures. Joint ventures might reduce competition (both in the R&D stage, and later in the production stage). However, any resulting loss to consumers must be balanced against the gain from accelerating the R&D process.

6. Conclusions.

Sunk costs play a central role in antitrust economics, but as we have seen, their measurement becomes more complicated when market conditions evolve unpredictably. In the presence of uncertainty, there is an opportunity cost of investing, rather than waiting for the arrival of new information about the likely returns from the investment. Firms have options to invest, and when they exercise those options, they give up the associated option value. That option value is also a sunk cost, and must be included along with the direct costs of machines and factories. I have tried to clarify the nature and role of this option value, how it depends on the degree of market uncertainty, and how it affects the measurement of sunk costs in antitrust analyses.

We have seen that when a firm can obtain information by waiting, the cost of a capital investment includes the opportunity cost of giving up the option to wait. This lost option value is sunk cost, and it can be quite large. On the one hand, this means that barriers to entry may be greater than one would otherwise think from using a conventional (but as we have seen, incorrect) measure of sunk cost. On the other hand, it means that conventional price-cost margins may greatly overstate market power, because much of the competition occurs prior to

entry (e.g., by developing a better product), rather than post-entry. In addition, the magnitude of the option value depends on the extent of uncertainty over future market conditions. In markets that evolve rapidly and unpredictability (e.g., many high-tech markets), option value is likely to be extremely important, and a crucial part of antitrust analysis. In more stable and predictable markets (e.g., food processing), it is likely to be less important, and could probably be ignored. In general, the extent of volatility is a basic structural feature of a market that plays an important role in determining the full sunk cost of entry, and thus the number of firms that can be expected to enter and compete.

We have also seen that there are situations where the firms in an industry learn from the investment experiences of each other. In this case, market failure can result: All of the firms may rationally prefer to wait, hoping that one of the other firms will invest first, with the result that no investment takes place. In such cases, a merger or joint venture might be warranted, even if such a merger or joint venture leads to an increase in ex post market power.

Although I did not address this point, there can also be situations where a firm can learn only by investing itself. An example is an R&D program undertaken by only a single firm. The cost, and even the feasibility, of completing the program are unknown at the outset, but as the firm invests (i.e., spends money on R&D), the uncertainty is gradually resolved, and the firm learns more and more about the ultimate cost and feasibility of completing the program. This creates an incentive to invest early, because investing yields information. It means that a firm might invest even though the conventionally measured NPV of investing is negative. Likewise, a firm might enter a market (e.g., an airline might start flying a new route) to learn more about the costs it is likely to incur and revenues it might earn, even though doing so seems uneconomical based on a conventional discounted cash flow analysis. At first glance, entry in such cases might be viewed as predatory, whereas in fact it is not.

Although I have focused on the antitrust implications of option value, it also has important implications for regulatory policy. Regulators, for example, typically ignore option value in rate-of-return regulation. In Pindyck (2004, 2005), I have shown that for telecommunications, failing to account for option value has resulted in inefficiently low regulated lease rates for unbundled network elements that incumbent firms must make available to entrants. Hausman and Myers (2002) have shown that the same problem exists in the regulation of U.S. railroads.

I have stressed that option value can be an important component of sunk cost, especially in markets that evolve rapidly and unpredictably. I have also stressed that ignoring option value can lead to substantial errors in the estimation of sunk costs, and thus incorrect conclusions when evaluating entry barriers or other factors affecting antitrust analyses. I have not, however, explained how option value should be calculated as a practical matter. Clearly, estimating this component of sunk cost can be more complicated than estimating and adding up the capital costs, start-up and marketing costs, etc., that comprise the direct component of sunk cost. Furthermore, the estimation of option value is typically industry- or firm-specific, and in antitrust contexts, must be tailored to the particular question that is at issue. Fortunately, some of the methods that have been developed for the analysis of financial options can be applied here, so that the estimation of option value is certainly feasible.[16] And as I have tried to show, ignoring option value is often not an option; difficult or not, it must be included as part of an antitrust analysis.

---

[16] Real options are often (but not always) simpler to evaluate than financial options, because the time element that is so important for financial options is often not present. (Financial options usually have lifetimes of less than a year, while real options are often perpetual, i.e., never expire.) Over the past decade, a variety of methods have been developed for the analysis of real options. See, for example, Copeland and Antikarov (2001).

References

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, August 1970, 488—500.

Bain, Joe S., *Barriers to New Competition*, Harvard University Press, 1956.

Baumol, William J., John C. Panzar, and Robert D. Willig, *Contestable Markets and the Theory of Industrial Structure*, Harcourt, Brace, Jovanovich, 1982.

Brealey, Richard A., and Stewart C. Myers, *Principles of Corporate Finance*, 6th Ed., McGraw-Hill, 1998.

Carlton, Dennis W., "Why Barriers to Entry are Barriers to Understanding," *American Economic Review*, May 2004, **94**,

Carlton, Dennis W., and Michael Waldman, "The Strategic Use of Tying to Preserve and Create Market Power in Evolving Industries," *RAND Journal of Economics*, Summer 2002, **33**, 194—220.

Copeland, Tom, and Vladimir Antikarov, *Real Options: A Practitioner's Guide*, New York: Texere, 2001.

Dixit, Avinash K. and Robert S. Pindyck, *Investment Under Uncertainty*, Princeton University Press, 1994.

Dixit, Avinash K. and Robert S. Pindyck, "The Options Approach to Capital Investment," *Harvard Business Review*, May-June 1995.

Hausman, Jerry A., and Stewart C. Myers, "Regulating the United States Railroads: The Effects of Sunk Costs and Asymmetric Risk," *Journal of Regulatory Economics*, 2002, **22**, 287—310.

Hull, John C., *Options, Futures, and Other Derivatives*, 3rd Edition, Prentice Hall, 1997.

McAfee, R. Preston, Hugo M. Mialon, and Michael A. Williams, "What Is a Barrier to Entry?" *American Economic Review*, May 2004, **94**, 461—465.

McDonald, Robert, *Derivatives Markets*, Prentice Hall, 2004.

McDonald, Robert, and Daniel L. Siegel, "The Value of Waiting to Invest," *Quarterly Journal of Economics*, Nov. 1986, **101**, 707—728.

Paddock, James L., Daniel R. Siegel, and James L. Smith, "Option Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases," *Quarterly Journal of Economics*, August 1988, **103**, 479—508,

Pindyck, Robert S. "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review*, December 1988, Vol. 79, pp. 969-985.

Pindyck, Robert S. "Irreversibility, Uncertainty, and Investment," *Journal of Economic Literature*, Sept. 1991, **29**, 1110—1148.

Pindyck, Robert S., "Investments of Uncertain Cost," *Journal of Financial Economics*, August 1993, **34**, 53—76.

Pindyck, Robert S. "Mandatory Unbundling and Irreversible Investment in Telecom Networks" NBER Working Paper No. 10287, January 2004.

Pindyck, Robert S., "Pricing Capital under Mandatory Unbundling and Facilities Sharing," NBER Working Paper No. 11225, March 2005.

Pindyck, Robert S., and Daniel L. Rubinfeld, *Microeconomics*, 6th Ed., Prentice Hall, 2005.

Schmalensee, Richard, "Sunk Costs and Antitrust Barriers to Entry," *American Economic Review*, May 2004, **94**.

Weitzman, Martin L., "Contestable Markets: An Uprising in the Theory of Industrial Organization: Comment," *American Economic Review*, 1983, **73**, 486—7.
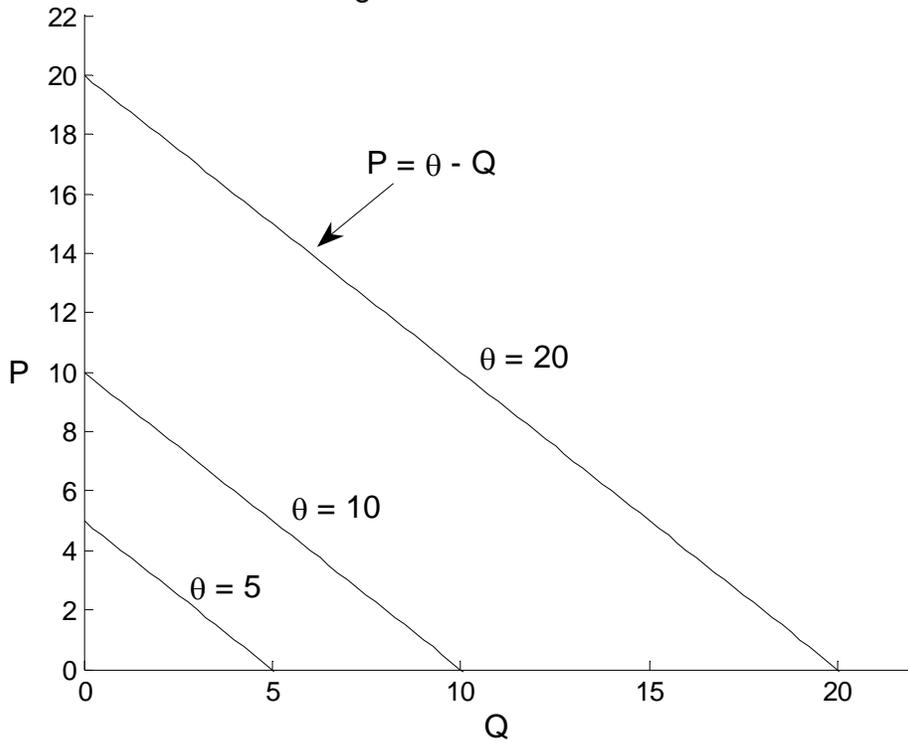
Figure 1: Demand Curve

Figure 1: Demand Curve

$P = \theta - Q$

$\theta = 20$

$\theta = 10$

$\theta = 5$

P

Q

Figure 2: Value of Option

Figure 2: Value of Option
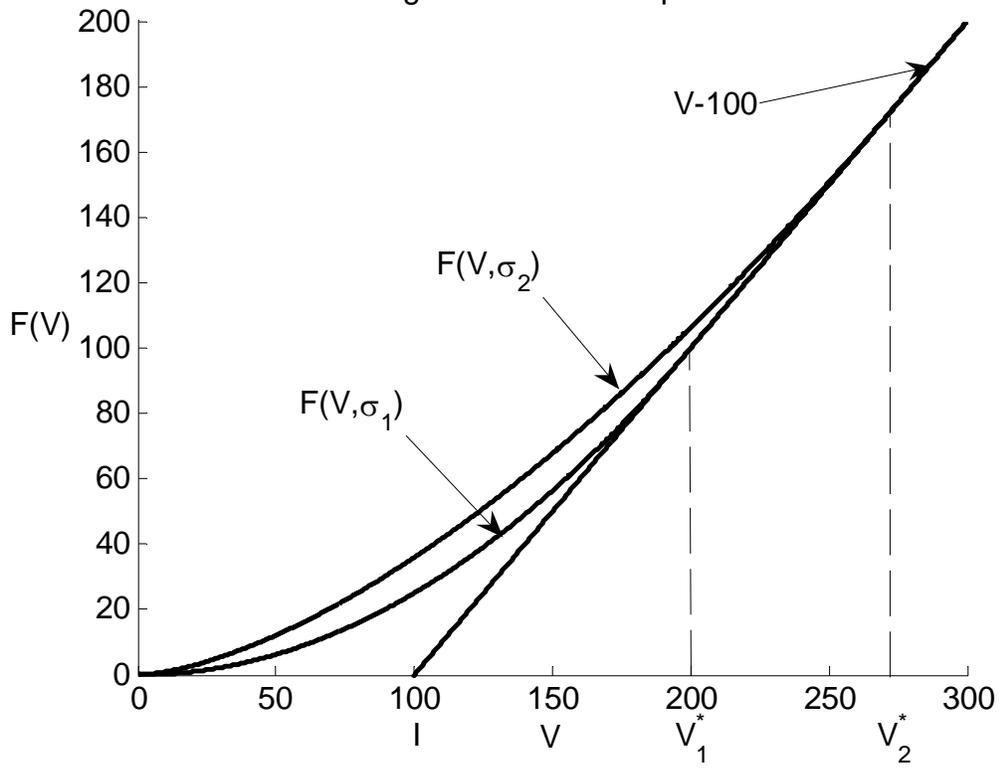
V-100

$F(V,\sigma_2)$

$F(V,\sigma_1)$

F(V)

I V V_1^* V_2^*

Figure 3: Increase in Sunk Cost, S_2/S