The Rental Cost of Sunk and Regulated Capital

John P. Small and Henry Ergas*

24 June 1999

Abstract

Standard approaches to estimating the cost of capital are vulnerable to two errors when applied to sunk assets subject to economic regulation. First, investors in sunk assets usually have a valuable ability to delay commitment which is therefore an opportunity cost of investment. Second, economic regulation alters the distribution of returns to capital, and may do so in a way that eliminates profit potential but leaves some risk of losses. This paper studies the impact of, and the interaction between, these effects. A fundamental connection is established between the value of the real option to delay investment, the risk of losses under regulation and the rate of economic depreciation. A practical benefit of this analysis is that existing empirical methods for estimating real options can now be used to estimate the size of the "investment incentive margin" that regulated firms should be allowed.

Address for Correspondence: John P. Small, Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand. *email:* jp.small@auckland.ac.nz

1. Introduction

This paper is concerned with the way that market risk and uncertainty interact with economic regulation, and the impact this has on the regulated firm. Regulators carry a heavy burden of responsibility for current and future outcomes in the markets under their control. As the interface between public policy, consumer interests and firms with market power, they are routinely required to make decisions with far reaching effects. Furthermore, in many jurisdictions it seems that regulatory activity is an increasing function of the number of deregulated markets opened to competition. This apparent paradox arises from the fact that the terms of access to bottleneck facilities is a crucial determinant of profitability for both entrants and incumbents in traditionally regulated industries such as telecommunications and electricity.

As the number of firms with an interest in regulatory decisions increases, there is a corresponding increase in the diversity and complexity of the arguments put to regulators. Two such arguments provide the primary motivation for this paper. Considerable attention has recently been focussed on the implications for regulation of the modern "real options" theory of investment, with views as to the relevance of this concept being sharply divided. In addition, firms have expressed concerns over the one-sided nature of their risk under regulation. By removing all of the "upside", the argument goes, regulation leaves the firm playing an unwinnable game.

Both of these arguments can be reduced to questions about how the opportunity cost of the capital employed by the regulated firm is affected by a particular one-sided random variable. The value of the real option to delay varies with the parameters of the firm's operating environment but can never be negative. Similarly, many common forms of regulation aim to prevent firms from securing economic rents but do not provide insurance against losses.

The primary goal of this paper is to study the connections between the delay option value and the expected cost of losses under regulation. Intuition suggests that these variables may share more than just the fact that they are both one-sided. As a trivial example, suppose that the firm expects that if it invests now, it will incur losses over the forthcoming period, but by waiting one period it will avoid such losses and break even for ever. In this case, the value of the real option to delay investment is exactly equal to the losses expected. It turns out that this relationship between real options and regulatory losses also holds in more general settings.

The ideas behind real options theory of investment can be traced back at least to Arrow (1968) but the theory itself is usually dated from McDonald and Siegel (1986). Pindyck (1991) offers an early survey, and Dixit and Pindyck (1994) provide a comprehensive and rigorous exposition of most aspects of the theory. More recently Abel et al (1996) have developed the linkages between this theory and earlier work on adjustment cost models. Despite the considerable breadth of the real options literature, there is very little work that considers the implications for regulated industries. The notable exceptions are Teisberg (1993) who studies the impact of regulatory risk on the size of investment projects undertaken, and Dixit and Pindyck (1994, chapter 9) who show how regulatory constraints can be embedded into a continuous time model of real option values. An interesting feature of this latter work is that it shows how regulation can be self-justifying. This occurs because the regulatory constraint reduces investment, thereby congesting the existing capacity, increasing the probability that the constraint will bind, and increasing the long-run average price to consumers.

While the issues studied here do not appear to have been considered previously, there are nevertheless some links to previous papers. Perhaps the most obvious connection is to the well known work of Averch and Johnson (1962). In that paper, the firm is assumed to be permitted a rate of return that exceeds its cost of capital, with the result that the firm makes excessive investment in unproductive assets. By refining the definition of the cost of capital for a regulated firm, our work could be interpreted as suggesting that the additional margin assumed by Averch and Johnson is not necessary. Provided that costs are defined correctly, all that is required is for the regulator to reimburse these costs.

This paper could also be thought of as part of a more general re-evaluation of the nature of risk and uncertainty. Pratt and Zeckhauser (1996), for example, cite "an explosion of literature addressing the severe errors society makes in confronting risks". Our analysis is consistent with this theme in the sense that it shows how the very act of regulating changes the nature and value of the risks faced by the regulated firm. Unless this endogeneity is recognised and accommodated, current costs may be under-recovered and efficient investment deferred.

Notwithstanding these connections, it may be most helpful to view this research as being a study of regulatory consistency. The regulatory process requires a determination of both the value of the firm's assets, and the cost of the capital invested in them. Assume that the firm's stock is freely traded under the expectation of normal (zero) profits only. If the issues studied here are relevant but neglected, then effect of regulation is to understate the cost of capital. This

disturbs the capital market equiulibrium, prompting investors to sell the stock and reducing the value of the firm to a level that is consistent with the regulated rate of return. In this case, there is a fundamental inconsistency between the regulated asset value and the permitted rate of return. Our work shows how to test for the existence of such an inconsistency, and to correct it if necessary.

The remainder of the paper is organised as follows. In section 2 we present a very simple model of real options and regulatory losses. When capital lasts for one period only, the option value and the one-sided costs of regulation are identical. While supportive of the intuition that motivated this work, section 2 leaves unanswered the question of how sunk assets affect the result. This is addressed in section 3 where we assume that capital is long-lived and that the regulator applies economic depreciation. This work confirms the results of the first model, showing that the real option to delay investment is identical to the expected cost of losses over the forthcoming period. In addition, it shows that with appropriate timing qualifications, economic depreciation is identical to the option value, and hence is also equal to the expected cost of losses. Option value models can therefore be used to estimate economic depreciation. Section 4 analyses the impact of regulation on CAPM based estimates of the risk premium. This emphasises the fact that the CAPM beta is in fact endogenous to the regulators decision and shows how the zero-profit regulatory constraint can be derived in a way that is consistent with this endogeneity. Section 5 presents some numerical simulations of the size of the option value mark-up based on the model derived in section 3. The paper closes with a summary of the results and their implications.

2. Options, Losses and Regulation

In this section we present a simple model of the impact of regulation on capital costs. We assume that the regulated firm has a particular investment project in mind but has not yet sunk any capital. The service life of capital invested is one period only. We assume that the regulator (a) values the asset base and (b) sets a maximum rate of return that can be earned on this base. The regulatory constraint may be expressed as a price which declines over time as would occur in price-cap regulation. Our results are readily extended to this case. The single essential feature of the type of regulation assumed here is that it is one-sided: the upside potential is limited but the downside is not.

We know from the real options literature that the firm will not invest unless two conditions are met. First, the project must be expected to cover its costs;

4

this is the standard NPV criterion. Secondly, it must not be the case that the firm expects to profit from delaying the investment. The value of the option to delay investment (henceforth refered to as just the option value) is the difference between the expected value of the firm if it invests now, and the expected value if it defers the investment decision until next period, which can be written as

$$\theta_t = I(R_{t+1} - R_t) \quad ; \quad t = 1, 2, ... \tag{2.1}$$

where R_t is the expected value of holding the additional asset in period t, expressed as a rate of return on capital invested. Similarly, R_{t+1} is the expected value of not holding the asset in period t but being free to invest at $t + 1$. The expectations embodied in R_t and R_{t+1} are both formed at time t, and the indicator function $I(x)$ has the property that $I(x) = x$ when $x > 0$ and $I(x) = 0$ otherwise.

If $R_t \geq R_{t+1}$ then $\theta = 0$ and the flexibility to delay investment has no value. In this case the firm will invest immediately provided R_t exceeds the cost of capital. Whenever $\theta_t > 0$, however, the firm does not invest because waiting until the next period has a higher expected return than investing immediately. Thus, θ_t is a one-sided random variable which varies with random shocks the the operating environment of the firm but can never be negative.

Observe that, since R_t is a rate of return, it can be decomposed into a known and constant risk adjusted cost of capital[1] r and an economic profit rate e_t as:

$$R_t = r + e_t \quad ; \quad t = 1, 2, ... \tag{2.2}$$

Suppose that there is some probability that the project will not cover its costs. We can represent the expected value of losses over period t (also expressed as a rate or fraction of capital invested) as

$$\lambda_t = I(r - R_t) \quad ; \quad t = 1, 2, ... \tag{2.3}$$

where the indicator function emphasises the fact that λ_t is a one-sided random variable. Notice that λ_t can be interpreted as the actuarially fair insurance premium associated with the risk of losses over the forthcoming period. If we now add R_{t+1} to the RHS of (2.3) and subtract its decomposition from (2.2), we obtain

---

[1]This is a naive measure of the cost of capital because it does not take into account the regulated and sunk nature of the asset. It is the measure that would arise from standard regulatory use of the CAPM combined with the relevant adjustments for taxes.

$$\begin{aligned}
\lambda_t &= I(r - R_t + R_{t+1} - r - e_{t+1}) \\
&= I(R_{t+1} - R_t - e_{t+1}) \qquad\qquad (2.4)
\end{aligned}$$

This shows how the expected cost of losses over the current period is related to the investment timing decision and can be interpreted in several ways. To focus on the investment timing issue, note that the larger is R_{t+1} relative to R_t, the more likely it is that the current project will make a loss.

Now consider the impact of regulation aimed at removing economic rents. The effect in (2.4) is to set $e_{t+1} = 0$ with the result that, under regulation

$$\lambda_t = I(R_{t+1} - R_t) = \theta_t \qquad\qquad (2.5)$$

We summarise this result in a proposition.

Proposition 2.1. *When costs are known and constant and capital lasts for one period, the real option value is also the actuarially fair insurance premium for the risk of losses imposed by economic regulation.*

The situation described above, while somewhat artificial, is nevertheless illustrative of an important point. The real option value signals information that can alter the profitability of the firm. This alteration is usually interpreted as showing how profits can be increased above the break-even level by careful attention to timing. The concept applies more generally, however, and could be used to value the change in losses that would occur through delaying investment, despite the fact that firms would not willingly invest at all if expected profits were negative.

In a regulated environment, economic profits are not available but losses can occur. While the impact of relatively small losses may be difficult to discern, the spectacular strandings of generation assets in the USA electricity industry lie at the other extreme. Stranding is a discrete event which occurs when losses of the type defined in (2.3) accumulate sufficiently and the prospect of offsetting profits is negligible.

If capital were not already sunk, a firm considering entering a regulated industry would evaluate its timing decision just as real options theory predicts. Given the constraints of regulation, however, the option value would signal investment at precisely the same time that the downside risk of losses was eliminated. Thus,

the option value and the loss-insurance premium are the same for regulated firms[2], as stated in the proposition.

2.1. Implications and Limitations

What does the above proposition imply for regulatory policy? The answer clearly depends on the extent to which the assumptions underlying (2.5) are applicable in any given situation, and we discuss this issue shortly. Some more fundamental remarks are also required, however. First, the proposition highlights the fact that decisions over new investment in regulated industries are subject to the same forces that prevail elsewhere. Second, it is clear that regulatory regimes which limit earnings to the risk adjusted cost of capital must also consider possibility of the firm incurring losses. Since the firm has no alternative but to self-insure against losses, the cost of this insurance is a genuine opportunity cost of capital. Finally, the real option to delay investment can provide a basis for valuing the actuarially fair premium for this insurance, at least in this simple context.

There are two main limitations of the approach underlying (2.5). First, while it is clear that future investment will not be forthcoming unless the allowed rate of return is adjusted to $(r + \lambda_t) = (r + \theta_t)$, this does not necessarily mean that the regulated rate of return on previously sunk capital should also include an "option premium". Secondly, the effect of relaxing the rather extreme assumption that capital lasts for one period is as yet unexplored. These limitations are addressed in the next section.

3. A More General Model

In this section we relax the assumption that capital lasts for only one period. Further, although we will discuss investment as if the firm retains the opportunity to delay its commitment, this is merely a convenient fiction that allows us to derive the real option value. In fact, the assets are already sunk and our task is to calculate the opportunity cost of the capital so employed. Consequently, there are no scale issues to consider here; the size of the plant is already determined. Finally, while the regulator continues to limit upside returns as previously, she must now also consider the treatment of depreciation.

---

[2]Insurance against losses is largely redundant in the absence of regulation because firms can offset these agains the rents earned during profitable trading periods.

We specify a production function for period t which relates output Q_t to inputs of capital K_t and labor L_t as follows:

$$Q_t = F(K_t, L_t)$$

where $F(.)$ is homogeneous[3] of degree one in K_t, and concave and non-decreasing in each input. At the beginning of each period the firm selects its preferred level of labor and installs additional capital equal to I_t with the result that the capital stock evolves according to

$$K_{t+1} = (1 - \delta_t)K_t + I_t$$

where δ_t is the rate of economic depreciation over period t. At the beginning of each period the firm selects I_t and L_t with the objective of maximising the following expected sum of discounted cash flows

$$E_t \sum_{t=1}^{\infty} r^{t-1}(p_t Q_t - w_t L_t - C(I_t))$$

where the expectation E_t is formed using the period t information set, which includes all magnitudes dated $t-1$ and earlier. In this equation, r is the firm's cost of capital (defined as in section 2), $p_t = p(Q_t)$ and w_t are the price and wage rates at time t respectively, and $C(I_t)$ is the cost of installing new capital. Future values of Q_t, w_t and q_t are assumed to have random components, possibly in addition to some deterministic trends. Using the homogeneity of the production function, the maximum profit attainable in period t can be written as

$$h(K_t, Q_t, w_t) = K_t(g_t - r) = K_t(g(Q_t, w_t) - r)$$

so that g_t can be interpreted as the greatest rate of operating profit obtainable at time t from the installed capital base K_t. Note that the total maximum economic profit rate is the difference between the operating profit rate g and the cost of capital, r. It is clear that g will vary randomly over time with shocks to demand and input prices; we assume however that since r is derived with specific reference to market risk, it does not vary with normal business conditions.

---

[3]This assumption is convenient rather than critical. It allows us to readily separate the quantity of capital invested from the cost of this capital.

For analytical convenience[4], we assume that the g_t's are independent draws from a normal distribution given by

$$g_t \sim NID(\mu, \sigma^2) \tag{3.1}$$

We will shortly impose the regulatory constraint on the firm by altering this distribution, but for now consider the value of this firm in the absence of regulation, which is given by

$$V_t = K_t \sum_{t=1}^{\infty} d^{t-1}(1-\delta)^{t-1}(g_t - r). \tag{3.2}$$

where $d = (1-r)$ is the firm's discount rate. Using standard results on infinite geometric sums and the normal distribution, (3.2) can be rewritten as

$$V_t \sim N(\nu_t, \sigma_t^2)$$

where

$$\nu_t = \frac{K_t(\mu - r)}{1 - d(1 - \delta)} \tag{3.3}$$

$$\sigma_t^2 = \frac{\sigma^2(\mu - r)^2}{1 - d^2(1 - \delta)^2} \tag{3.4}$$

The impact of the homogeneity assumption can be seen in (3.3), where the mean value of the firm is linear in the capital base. While this would be an unreasonable approach if we were considering large changes to K_t, it is largely innocuous here since the capital base is already installed.

3.1. The Cost of Regulation

In this section we introduce the regulatory constraint by changing the distribution of the per-period maximum operating profit rate g_t. Define the regulated rate, g_t^R as:

$$g_t^R = \begin{cases} g_t & \text{if} \quad g_t < \alpha \\ \alpha & \text{if} \quad g_t \geq \alpha \end{cases} \tag{3.5}$$

---

[4]It is straightforward to derive our results under more general distributional assumptions, including non-stationary elements such as a time dependent mean.

Thus, g_t^R is a censored[5] random variable which depends on the underlying latent variable g_t, defined in (3.1). The regulated return variable g_t^R takes the value α during good trading periods (when economic profits would have been made in the absence of regulation) and g_t in all other periods. The point α at which the regulatory constraint binds is left unspecified for now.

Using standard results on the mean of a censored normal distributions[6], the expected rate of operating profit $E(g^R)$ can be written as

$$E(g^R) = \mu^R = \alpha - (\alpha - \mu)\Phi(\frac{\alpha - \mu}{\sigma}) - \sigma\phi(\frac{\alpha - \mu}{\sigma}) \tag{3.6}$$

where $\phi(.)$ and $\Phi(.)$ denote the standard normal density function and CDF respectively. This allows us to write the expected economic rate of return on regulated capital as

$$E(g^R - r) = \mu^R - r \tag{3.7}$$

As discussed further below, the variance of g_t^R also differs from that of g_t. Concentrating on the means, however, observe that the difference between the value of the regulated and unregulated firms can be found by substituting μ_t^R for μ in (3.3) and subtracting the result from ν_t to get

$$
\begin{aligned}
\Delta^R E(V_t) &= \frac{K_t(\mu - \mu^R)}{1 - d(1 - \delta)} \\
&= \frac{K_t}{1 - d(1 - \delta)}\left(\sigma\phi(\frac{\alpha - \mu}{\sigma}) + (\alpha - \mu)(\Phi(\frac{\alpha - \mu}{\sigma}) - 1)\right) \quad (3.8)
\end{aligned}
$$

The mean difference as expressed in (3.8) follows directly from the specification of the censored distribution in (3.5).

When regulation is imposed, therefore, the value of the regulated firm is reduced by an amount that can be estimated using (3.8). Such a reduction is, of course, merely a natural consequence of restraining the firm from exercising its market power and is not, of itself, cause for concern. Since regulation removes the upside earning power of the firm, however, there is a danger that the continued

---

[5]There is a subtle but important difference between "censoring" and "truncating" a sample of data. In the former case, all observations are present but the true values of some are incorrectly reported as taking a limiting value. Truncation refers to a situation in which some values are simply not reported.

[6]See, for example, Greene (1990) pp725-7.

presence of downside risk may lead to economic losses on average. Note that the expected economic loss that arises as a result of regulation can be found using (3.7). As in section 2, we use λ_t to represent these losses, which can be expressed as:

$$\lambda_t = I(r - g_t^R) \tag{3.9}$$

where the indicator function $I(.)$ is as defined for (2.1) above.

3.2. The Real Option to Delay Investment

Having derived the expected cost of losses under regulation, we now consider the option value in this model. Retaining the focus on sunk assets, we deliberately adopt the fictitious assumption that these assets have not been sunk. We consider, therefore, the decision of a firm which is contemplating the investment of K_t at time t, and realises that delaying this investment may be valuable. The firm depicted here is regulated but it should be clear that the analysis is exactly the same for an unregulated firm except that the operating profits would be denoted g_t rather than g_t^R in this case.

It will be convenient to split the economic profits in (3.2) into those earned in period t, which we denote R_t and those from all subsequent periods, as follows

$$
\begin{align}
R_t &= E_t(g_t^R - r) \tag{3.10} \\
R_{t+} &= E_t \sum_{t=2}^{\infty} r^{t-1}(1-\delta)^{t-1}(g_t^R - r) \tag{3.11}
\end{align}
$$

The timing decision of the firm depends heavily on the value of current period cashflows which are only available if the investment occurs at time t. For future reference, note R_t is just the same as the expression given in (3.7).

3.2.1. Economic Depreciation

Depreciation is also an important aspect of the firm's timing decision. However, depreciation in periods beyond t are not relevant when the firm decides whether to invest now or wait one period; it is peroid t depreciation that matters for this decision. It is well known that the economic depreciation of capital arises from the erosion of two stocks: total willingness to pay for the services of the asset, and total service life of the asset. Reductions in either or both of these stocks reduce

11

the value of the firm holding the asset, and we can therefore define the *expected* rate of depreciation of capital over period t reference to the expected value of the net cashflows.

$$\delta_t = E_t \frac{V_t - V_{t+1}}{V_t}$$

which can be rewritten by substituting (3.10) and (3.11) into (3.2) as follows

$$
\begin{aligned}
\delta_t &= \frac{K_t(R_t + R_{t+}) - K_t R_t}{K_t(R_t + R_t)} \\
&= \frac{R_t}{R_t + R_{t+}}
\end{aligned}
\tag{3.12}
$$

This shows the relationship between economic depreciation over the period and the share of total lifetime discounted net cash-flows that the asset delivers over the period. The use of economic depreciation is important here because, like the real option value it is a one-sided random variable that depends on the value of the firm in successive time periods.

Note, however, that δ_t is a forward looking measure of depreciation that has not yet occurred. In what follows we assume that the regulator correctly anticipates δ_t and uses this measure when calculating the cost of capital.

3.2.2. The Option Value Calculation

We now derive the value of the option to delay investment. In doing so, we assume that the scale of the project would not change as a result of delay. The value of the unregulated firm when investment occurs at time t is given by (3.2) and substituting g_t^R for g_t in this expression gives the corresponding value for the regulated firm. If, on the other hand, investment is delayed until period $t+1$, the value of the firm is anticipated (at the beginning of period t) to be

$$
\begin{aligned}
E(V_{t+1}) &= \frac{K_t}{1 - \delta_t} \sum_{t=2}^{\infty} d^{t-1}(1 - \delta)^{t-1} g_t^R. \\
&= \frac{K_t}{1 - \delta_t} R_{t+}
\end{aligned}
\tag{3.13}
$$

12

where it will be observed that the capital stock is adjusted upwards to account for the lack of depreciation over period t when investment is delayed until period $t + 1$. Using (3.12) we can rewrite this as

$$
\begin{aligned}
E(V_{t+1}) &= K_t \frac{R_t + R_{t+}}{R_{t+}} R_{t+} \\
&= K_t(R_t + R_{t+})
\end{aligned}
$$

We can now derive the real option value as the difference between the expected value of the firm from investing in the current and future periods.

$$
\begin{aligned}
K_t \theta_t &= I(E(V_{t+1}) - E(V_t)) \\
&= I(K_t [R_t + R_{t+} - (R_t + R_{t+})]) \\
&= 0
\end{aligned}
$$

Thus, when future economic depreciation is correctly estimated, there is no real option value. Firms are indifferent between investing now and delaying one period because economic depreciation refunds any loss in value that occurs as a consequence of investing too early.

Furthermore, if no allowance is made for deprecation, then the option value is identical to the expected cost of losses, as was the case in the simple model of section 2. To see this, note that with no allowance for depreciation, (3.13) would become

$$
E_{NoDep}(V_{t+1}) = K_t R_{t+}
$$

in which case the option value would be

$$
\begin{aligned}
K_t \theta_t &= I(-R_t) \\
&= I(r - g_t^R).
\end{aligned}
$$

the right hand side of which is identical to that of equation (3.9). We summarise these results as follows.

Proposition 3.1. *The expected cost of losses for a regulated firm is the same as the real option value that would apply if the assets were not sunk. This is also equal to the rate of economic depreciation that is expected to occur over the forthcoming period.*

This result confirms the intuition developed in the simple model presented at the start of this paper. There is a fundamental duality between the one-sided risk of making a loss on sunk and regulated capital and the value of the real option to delay investment in this capital. Indeed, these are seen to be simply opposite sides of the same gamble.

3.3. Interpretation

This analysis has some important practical implications and can be interpreted in several ways. First, the connection between option values and expected losses helps to clarify the true opportunity cost of capital sunk in regulated assets. Unless downside insurance is provided in some way, the elimination of economic profits will erode the value of the firm *below the level assumed by the regulator*. The issue can therefore be interpreted as concerning the internal consistency of the rate setting process. It is easy enough to set the expected value of profits to zero, but unless the expected cost of losses is also zero, the regulator has used an inconsistent model to set the rate of return on capital.

Secondly, if economic depreciation is correctly applied, the expected cost of these losses is zero.

Thirdly, the real option to delay investment in the sunk assets, which would be directly relevant if investment had not already occurred, is identical to the rate of economic depreciation. This opens up the possibility of using the well established real options empirical methodology to derives estimates of economic depreciation. Since real options models typically impose exogenous depreciation rates, it appears that any remaining option values arise from the difference between the assumed exogenous rate of depreciation and the expected rate of economic depreciation. A further benefit of this connection is that a test of the hypothesis that the real option is zero can be interpreted as a test of the consistency of the rate setting model.

Finally, the derivation of real options for regulated firms leads to a clear answer the question of how much "profit" a regulated firm should be permitted in order to provide an incentive to make efficient investments. Our analysis suggests that efficient investment will occur in regulated firms provided that costs are measured carefully, although there is reason to believe that regulated firms will invest less than unregulated firms (Small, (1999)).

14

4. Regulation and the CAPM

In this section we consider the same problem from a different perspective. As in the previous section, we assume that demand is random but the price (or rate of return) is fixed by the regulator at a level that is consistent with zero expected economic profits over the lifetime of the asset being regulated. The price includes an allowance for economic depreciation and a risk premium estimated by the CAPM.

It is well known that a CAPM based risk premium only rewards investors for so-called systematic risk, which increases the risk of the market portfolio. When regulators use the CAPM as a basis for estimating the cost of capital, the resulting risk adjusted rate of return depends on the covariance of firm returns with the return on the market portfolio (σ_{iM}/σ_M^2), the market risk premium $(r_M - r_f)$ and the risk free rate of return r_f as follows:

$$r_i = r_f + (r_M - r_f)\beta_i \tag{4.1}$$

where

$$\beta_i = \frac{\sigma_{iM}}{\sigma_M^2} \tag{4.2}$$

Assume that both r_i and r_M are random variables and consider the impact on β_i of censoring the distribution of r_i in a manner that mimics the effect of regulation. To model this, we use the specification defined in the previous section and assume that the censored return series r_{it} and the original return series r_i are related as

$$r_{it} = \begin{cases} r_i & \text{when } r_i \leq \alpha \\ \alpha & \text{when } r_i > \alpha \end{cases}$$

This specification has the effect of censoring the distribution of r_i such that any returns above a specified level α are restricted to α. The following proposition describes the relationship between the covariance of r_i with r_M and the covariance of r_{it} with r_M.

Proposition 4.1. *If firm i is small relative to the market, censoring of the right-hand tail of r_i reduces the absolute value of σ_{iM} and the size of the reduction is given by*

$$\sigma_{itM} - \sigma_{iM} = E(r_M - \mu_M) \left(\int_\alpha^\infty (r_i - \alpha) f_r(r) dr - I(r_i - \alpha) \right) \qquad (4.3)$$

where $f_r(r)$ is the density of the unrestricted returns distribution and I is an indicator variable that takes the value 1 when $r_i > \alpha$ and zero otherwise.

The proof of this result is presented in the appendix. Note, however, that the final bracketted term in (4.3) is the difference between two terms that differ only in respect of the weighting applied to them. Since the weights on the first must be less than unity and the second term is unweighted, this difference is negative and $\sigma_{itM} < \sigma_{iM}$. Thus, regulation reduces the numerator for the regulated firm's CAPM derived beta.

4.1. The Implications of Censoring

If the regulated firm comprises a large part of the market portfolio, then regulation will also reduce σ_M^2. For most practical cases, however, it seems reasonable to assume that the actual and mean returns on the market portfolio, and the variance of these, does not change as a result of firm i being regulated. In this case, regulation has two effects on the CAPM. First, regulation reduces the risk premium that would be predicted by valid application of the CAPM. This is clearly evident from (4.3) and the discussion following it and highlights a further justification for the relevance of the issues studied here. Since the firm's beta is endogenous to the regulator's decision making process, failure to recognise this will (except by sheer chance) result in an inconsistent estimate of the cost of capital

Secondly, regulation makes a valid application of the CAPM more difficult. In particular, it invalidates the standard practice of estimating the CAPM beta using least squares regression because the dependent variable in the regression is censored[7]. This effect can be quite significant and is likely to become more important, the more frequently the regulatory constraint binds. Interestingly, however, this bias *increases* the estimated size of beta and therefore tends to over-compensate the firm for systematic risk.

---

[7] It is well known in the econometrics literature that OLS regression estimates of the slope parameter are biased and inconsistent when the dependent variable is censored; see Davidson and MacKinnon (1993), for example, or any other good econometrics text.

4.1.1. Simulation Evidence on Beta Bias

To investigate the empirical significance of bias in OLS estimates of beta this effect, a small simulation experiment conducted using the SHAZAM (1993) software. The experiment used a sample size of 100 and the number of replications for each case was 1000. The cases differed by the true (unregulated) value of beta and the percentage of observations that were censored, which we interpret as the percentage of time periods in which the regulatory constrain is binding. The same random numbers were used for each case, so the only differences are those attributable to the parameters reported.

Since valid estimates are obtained using a Tobit model, our experiment focussed on the *difference* between tobit and OLS estimates of beta. In table 1, we report the mean difference across the 1000 replications[8], expressed as $\widehat{\beta}_{Tobit} - \widehat{\beta}_{OLS}$. The predominately negative numbers in this Table show that OLS overstates beta for regulated firms in general.

Table 1: Bias in OLS Estimates of Beta for Regulated Firms

Censoring %	$\beta = 0.5$	$\beta = 1.0$	$\beta = 1.5$
0	-0.004	-0.035	-0.100
10	-0.050	-0.129	-0.259
20	-0.043	-0.067	-0.097
30	-0.038	-0.032	0.001
40	-0.051	-0.063	-0.039
50	-0.094	-0.156	-0.204
60	-0.151	-0.292	-0.458
70	-0.225	-0.474	-0.759
80	-0.330	-0.709	-1.123
90	-0.477	-1.043	-1.600

This small simulation experiment demonstrates the importance of regulatory censoring on the cost of capital. Since the size of the bias that applies in any given case cannot be known *a priori*, it is important the estimates of beta for regulated firms are obtained using statistically methods. Furthermore, despite the fact that the bias in OLS estimation of regulatory betas tends to offset the impact of censoring on the risk of losses, the only way to know which of these effects dominates is to calculate each in a consistent manner.

---

[8]We specified the market risk premium as being equal to 5 plus the sine of the observation number $t = 1, ..., 100$. The firm return was formed by summing a risk free rate (also 5), the market risk premium multiplied by the unregulated beta, and a standard normal error term.

4.2. Adjusting the Regulated Rate of Return for Diversifiable Risk

Since beta is endogenous, the solution to the regulator's rate setting problem must be solved with reference to the underlying variables. A convenient way to think of the task at hand is to ask the following question. What α should be selected if the aim of the regulator is to reduce the expected nominal rate of return to a level just consistent with the risk borne by investors? If we can find such a value for α then the regulator can (at least in theory) prevent the firm from earning economic rents without imposing losses, on average.

This problem is solved by finding the α which results in $\mu_{it} = 0$. Using equation (8.1) from the appendix, the relevant condition can be written as

$$\mu_i - \int_\alpha^\infty r_i f_r(r) dr + \alpha \int_\alpha^\infty f_r(r) dr = 0$$

from which we can deduce that the appropriate truncation point is given by

$$\alpha^* = \frac{-\int_{-\infty}^\alpha r_i f_r(r) dr}{1 - \int_{-\infty}^\alpha f_r(r) dr} \tag{4.4}$$

This expression can be interpreted as showing what the limit on the regulated firm's returns should be if the firm is to make zero economic profits overall. Thus, one way to estimate the regulated rate of return is to substitute (4.4) and (4.3) into (4.2) to derive the correct measure of diversifiable risk and then set the regulated rate with reference to this, taking account of taxes and other effects in the usual manner. By embedding the effect of regulation into the rate setting process, this approach results in an internally consistent solution to the regulator's problem. It is, however, a difficult and complicated solution relative to the option value mark-up proposed in the previous section. Moreover, since this method is designed to solve the same problem as that addressed in section 3, the end results of these two approaches should be identical.

5. The Size of Option Value Adjustments

Given the difficulties of setting the regulated rate through the adjustments de-fined in section 4, it seems preferable to use the option value approach described in section 3. This method begins with a valid estimate of the regulated firm's CAPM beta, which is combined with the usual adjustments for taxes to form what is frequently referred to in regulatory circles as the "weighted average cost

of capital", or WACC. This rate can then be adjusted upwards by an amount equal to the expected cost of losses under regulation. The size of the adjustment is exactly equal to the real option value associated with delaying investment in the currently sunk assets for one period, and can be estimated by whatever method is relevant, given the nature of the business concerned.

If unregulated market returns for the business under study can reasonably be represented by a normal distribution, then the parametric models of section 3 can be used to estimate the size of the adjustment required. We can define the adjustment as follows:

$$\Delta = (\alpha - r)|_{\mu^R = r}$$

where α is the final (adjusted) regulated rate of return, r is the WACC, and μ^R is defined by (3.6). Now, by substituting in the conditioning event $\mu^R = r$, the mark-up that is required over r can be written as

$$\Delta = \alpha - \mu^R|_{\mu^R = r}$$

from which it can be seen that

$$\Delta = \left[(\alpha - \mu)\Phi(\frac{\alpha - \mu}{\sigma}) + \sigma\phi(\frac{\alpha - \mu}{\sigma}) \right]|_{\mu^R = r}.$$

Given the parameters μ and σ of the unregulated distribution of returns, the size of an estimate of Δ can be readily computed for any α. This estimate can be refined sequentially by making small changes to α until the condition $\mu^R = r$ is satisfied. Figure 1 presents the results of a number of such processes in graphical form. It shows that as the mean value of the unregulated distribution of returns (μ) increases, the size of the adjustment (Δ) required to account of censoring this distribution through regulation, falls. This makes intuitive sense: the more inherently profitable is the firm, the larger is μ and the smaller is the probability of it even not covering its costs; consequently, the regulated rate of return only has to exceed r by a small amount.

Figure 1 also shows that the more risky are the firm's profits, the greater is the adjustment required, other things being equal. This is further corroboration of the link between real options (which are known to be increasing with the variance of returns) and the risk of losses under regulation.

The size of the adjustment required is an empirical matter to be assessed on a case by case basis. Particular care needs to be taken to specify the industry

19

model carefully. Dixit and Pindyck (1994) present several prototype models which permit "off the shelf" methods to be used in many cases, particularly those in which demand can be reasonably represented as a geometric Brownian motion. This approach will generally require the estimation of key demand parameters (the growth rate and its variability) which are used to construct the real option value[9].

6. Conclusion

This paper was motivated by the problems currently faced by regulators in assessing the merits of two particular arguments used in the rate setting process. It is certainly conceivable that regulated firms could incur losses as a result of regulation. Similarly, there is little doubt that firms behave as if the real option theory of investment were correct. There is, however, a natural scepticism about such arguments since they are quite new (this is particularly true of the option value approach) and would have the effect of increasing the permitted rate of return if they were to be accepted. Also, in respect of the option value argument, it is frequently asserted that this effect is adequately captured by using economic depreciation rather than some arbitrary rule.

Our analysis shows that there is a fundamental connection between the expected cost of losses arising from the one-sided nature of economic regulation, and the real option to delay investment. Indeed, if no allowances are made for depreciation, these are in fact the same thing. If the assets were not already sunk, the regulated firm would only choose to hold them until the next period if no losses were anticipated during this time. Otherwise, the firm would rationally prefer to delay the investment decision for one period. The value of the option to delay investment is just the difference between the value of these alternatives. Hence, if losses are expected over the current period, the real option to delay investment is exactly equal to the value of these losses.

We have also shown, however, that the amount of economic depreciation expected over then forthcoming period is identical to the value of the real option to delay investment over this period. Consequently, empirical models for estimating

---

[9]In unpublished work, the authors have applied this approach to regulated industries in Australia. Our estimates correspond to an estimate for Δ of around twenty percent of the cost of capital (r) for both telecommunications and rail transport. Note that the option values from Brownian motion models are applied multiplicatively to r, rather than additively as in the formulations presented in this paper.

real options can be used to estimate economic depreciation on a forward looking basis. If regulated rates of return are adjusted using real option values, the expected value of the risk of losses is zero and firms have an adequate incentive to invest.

7. References

Abel, A.B., A.K. Dixit, J.C. Eberly and R.S. Pindyck (1996) "Options, the value of capital and investment", *Quarterly Journal of Economics*, 111, pp. 753-777.

Arrow, K.J. (1968) "Optimal capital policy with irreversible investment" in J.N. Wolfe (ed.) Value Capital and Growth, Essays in Honour of Sir John Hicks, *Edinburgh University Press*, Edinburgh.

Averch, H. and L. Johnson (1962) "Behaviour of the firm under regulatory constraint", *American Economic Review*, 52, p1052

Davidson, R. and J.G. Mackinnon (1994) Estimation and Inference in Econometrics, *Oxford University Press*: New York.

Dixit, A.K. and R.S. Pindyck (1994) Investment Under Uncertainty, *Princeton University Press*: Princeton: New Jersey.

McDonald, R. and D. Siegel (1986) "The value of waiting to invest", *Quarterly Journal of Economics*, 101, pp. 707-728.

Pindyck, R.S. (1991) "Irreversibility, uncertainty and investment", *Journal of Economic Literature*, 29, 1110-1152.

Pratt, J.W. and R.J. Zeckauser (1996) Willingness to pay and the distribution of risk and wealth, *Journal of Political Economy*, 104, pp.747-763.

SHAZAM (1993) Users Reference Manual, McGraw Hill: New York.

Small, J.P. (1999) The timing and scale of investment under uncertainty, CRNEC Discussion Paper, University of Auckland.

Teisberg, E.O. (1993) "Capital investment strategies under uncertain regulation", *Rand Journal of Economics*, 24, pp. 591-604.

8. Appendix

Let μ_i and μ_{iC} denote the mean of the unrestricted and restricted distributions of returns respectively. The relationship between these is

$$
\begin{aligned}
\mu_{iC} &= \int_{-\infty}^{\alpha} r_i f_r(r) dr + \alpha \int_{\alpha}^{\infty} f_r(r) dr \\
&= \int_{-\infty}^{\infty} r_i f_r(r) dr - \int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr \\
&= \mu_i - \int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr \qquad\qquad (8.1)
\end{aligned}
$$

Note that since $r_i > \alpha$ everywhere in the censored section of the distribution, $\mu_{it} < \mu_i$. Now observe that the censored return series r_{it} can be written as

$$
r_{iC} = r_i - I(r_i - \alpha)
$$

where $I = 1$ when $r_i > \alpha$ and $I = 0$ otherwise. Combining the last two equations, we have

$$
r_{iC} - \mu_{iC} = r_i - \mu_i + \int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr - I(r_i - \alpha)
$$

from which, using the assumption that firm i is small and hence trucation does not affect the market outcome, we deduce that

$$
\begin{aligned}
\sigma_{iCM} &= E(r_M - \mu_M)(r_{iC} - \mu_{iC}) \\
&= E(r_M - \mu_M) \left[(r_i - \mu_i) + \int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr - I(r_i - \alpha) \right] \\
&= \sigma_{iM} + E(r_M - \mu_M) \left[\int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr - I(r_i - \alpha) \right]
\end{aligned}
$$

This establishes the validity of the expression in the Proposition. To infer that the absolute value of σ_{iCM} is less than that of σ_{iM} we observe that

$$
\int_{\alpha}^{\infty} (r_i - \alpha) f_r(r) dr < I(r_i - \alpha)
$$

since these are essentially the same term apart from the density weights which reduce the size of the left hand side. Thus, the expected value of their difference

is always negative. If r_i covaries positively with r_M then observations in the right-hand tail of r_i will occur at the same time as observations in the right-hand tail of r_M. In this case $(r_M - \mu_M)$ is postive and $\sigma_{iCM} < \sigma_{iM}$. If r_i covaries negatively with r_M then observations in the right-hand tail of r_i will occur at the same time as observations in the left-hand tail of r_M. In this case, both σ_{iM} and $(r_M - \mu_M)$ are negative, but σ_{iCM} is less negative than σ_{iM}. Q.E.D.
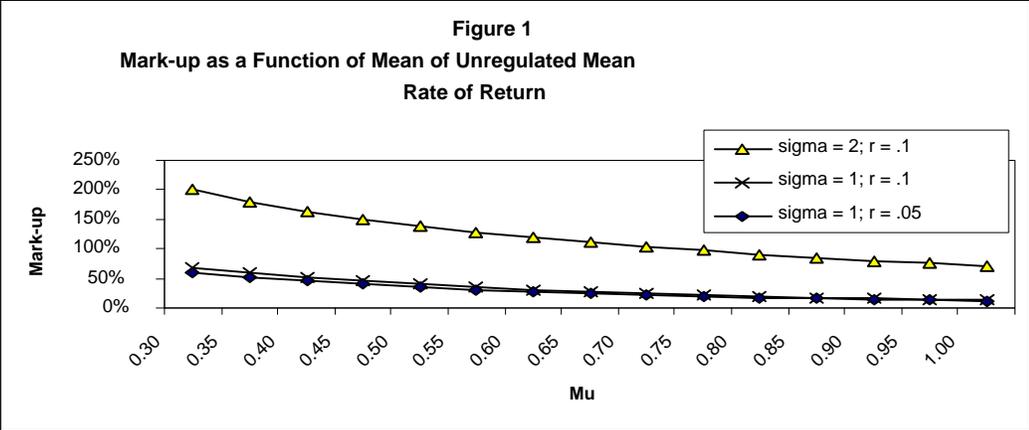
Figure 8.1: