Risk, Price Regulation, and Irreversible Investment

Lewis T. Evans*

New Zealand Institute for the Study of Competition and Regulation, and Victoria University of Wellington, Wellington, New Zealand

Graeme A. Guthrie

Victoria University of Wellington, Wellington, New Zealand

December 28, 2004

*Corresponding author. Address: ISCR, PO Box 600, Victoria University of Wellington, Wellington, New Zealand. Ph: 64-4-4635562. Fax: 64-4-4635566. Email: lew.evans@vuw.ac.nz

Risk, Price Regulation, and Irreversible Investment

Abstract

We show that regulators' price-setting, rate base, and allowed rate of return decisions are inextricably linked if prices are set so that regulated firms just break even whenever they are forced to invest. Breaking even ex ante is a necessary condition for Ramsey pricing to be sustainable over time. Unless regulators adopt traditional rate of return regulation, the irreversibility of much infrastructure investment significantly alters the results of the approach to price-setting described by Marshall, Yawitz and Greenberg (1981). In particular, the practice of 'optimizing' inefficient assets out of the regulated firm's rate base, as occurs in total element long-run incremental cost calculations in telecommunications, exposes the firm to demand risk. The firm requires an economically-significant premium for bearing this risk, and this premium is a function of both the systematic and unsystematic risk of demand shocks. In addition, we argue that if the firm is to break even under incentive regulation then the level of the rate base will exceed the optimized replacement cost by an amount which we interpret as the value of the excess capacity of the firm's assets. If this component is excluded from the rate base, incentive regulation will not be sustainable.

JEL Classification code: G31, L5

Keywords: Regulation, Cost of capital, Rate base, Sunk costs

Risk, Price Regulation, and Irreversible Investment

1 Introduction

A regulator needs to make three decisions when setting the prices which a regulated utility may charge. It needs to choose the appropriate cost of the firm's assets (the rate base), the rate of return the firm is allowed to earn on this rate base, and the prices the firm is allowed to charge. Marshall, Yawitz and Greenberg (1981) show how the last two decisions are inter-related. However, because they focus on traditional rate of return regulation, they do not discuss the effect of the choice of rate base. Nor do they consider the implications of irreversible investment, which characterizes most industries subject to price regulation.[1] In this paper we show that the choice of rate base can have a crucial impact on the other two decisions, and that the reason for this is the irreversibility of investment. In particular, we demonstrate that the regulator's choice of rate base and the form of regulation impact on the risks which the regulated firm faces, and thus on the rate of return it should be allowed to earn.

The Ramsey-pricing process for regulating a natural monopoly firm is, in a static setting, second-best welfare optimal in that it maximizes welfare subject to a zero profit constraint. We show how to implement the zero profit condition in a dynamic setting where the regulated firm is forced to make irreversible investments in order to meet demand. If welfare is to be maximized and the firm is to be financially viable in the long run, then the firm must just be able to cover the anticipated cost of worthwhile investments on a forward-looking basis.[2] We find, for a variety of forms of regulation, the rate of return that achieves this goal; our approach admits a mixture of markets supplied and products offered by the regulated firm. The way in which the zero profit condition is implemented is determined in large part by the regulator's choice of rate base, and this choice will affect the sharing of risk between consumers and producers, and thereby the overall level of welfare.[3]

There are two widely-applied rate bases. Traditional rate of return regulation uses the depreciated historical installation cost of existing assets as the rate base. When combined with

[1] Irreversibility is a widespread phenomenon, even in industries where physical capital is not especially industry-specific. For example, between 50 and 80 percent of the cost of machine tools in Sweden is sunk (Asplund, 2000), and the market value of physical capital in the US aerospace industry is just 28 percent of its replacement cost on average (Ramey and Shapiro, 2001). Irreversibility is likely to be even greater in most infrastructure networks. Hausman (1999) and Economides (1999) debate the extent of irreversibility in the context of telecommunications.

[2] Hausman and Myers (2002) argue that the provision of inadequate revenue to the regulated owner of the railroad infrastructure in the UK, Railtrack, led to under-investment in maintenance that contributed to accidents on that railway. Unable to raise funds in financial markets to finance investment in track upgrades, Railtrack was placed in liquidation by the UK government in October 2001.

[3] Cowan (2003) examines the effects of regulation in allocating price risk between consumers and shareholders, finding a trade-off between risk allocation and allocative efficiency.

forecast operating expenditure, this yields the revenue requirement that, along with forecast demand, is used to set prices.[4] The historical cost rate base continues to be used in some situations, including elements of the electricity transmission system in the US. Historical cost rate of return regulation was widely used until the 1980s when it was gradually replaced with incentive regulation, where prices are set in ways that seek to mimic competitive markets. A common approach is to periodically set prices using a rate base that is calculated using the least cost bundle of assets required to service existing customers, known as optimized replacement cost (ORC).[5] Newbery (1999, Chapter 7) argues that in telecommunications, where ORC is the total element long-run incremental cost (TELRIC) of a service, use of an ORC rate base will yield price paths that approximate, as well as is possible, those of competitive markets.[6] The TELRIC approach in telecommunications has been applied widely in the UK and the US, and is recommended by the European Commission (Newbery, 1999, p. 339).

We introduce a new rate base, optimized deprival value (ODV), that measures the cost to the firm if it is deprived of its assets. ODV exceeds ORC by an amount equal to the present value of expected cost savings resulting from the firm's 'excess' capacity. Our work suggests that use of ODV is necessary if the firm's allowed revenue is to be determined solely by the cost structure of a hypothetical efficient replacement firm; that is, the rate base should be ODV, and not ORC, in order for incentive regulation to be credibly implemented.

The revenue which the regulated firm requires if it is to break even equals the sum of expected operating costs, expected economic depreciation and a reasonable rate of return earned on the rate base. We show that if the regulator imposes a historical cost rate base, then the only risk that the firm must bear is the risk that demand and operating cost experience shocks after the price-setting process is complete (and before prices are reset in the future). If, instead, the regulator adopts ORC or ODV as the rate base, then the firm is also exposed to the risk of capital price shocks and, because fluctuations in demand affect the capacity of the (hypothetical) assets on which the optimization calculation is based, to the risk of future demand shocks at the time prices are reset.

Despite the fact that our analysis is performed with the Capital Asset Pricing Model (CAPM) as our valuation model, the irreversible nature of investment means that unsystematic demand

[4] Typically, approved investment plans also affect the revenue requirement under rate of return regulation. For a discussion of the process of rate of return regulation see Spulber (1989, pp. 270–279).

[5] An alternative approach, RPI-X regulation, involves setting a price path by allowing a starting price to grow at the rate of inflation, adjusted for industry-specific factors, such as relative productivity growth and input price changes, that are taken to be exogenous to the firm. The issues dealt with in this paper are directly relevant to the choice of X and the starting price, and to evaluating violations of such a regime. If the regulated firm is to just cover the costs of its investments, the factors of irreversibility, growth in demand, supply and capital prices that are important to these investigations will enter the analysis as described in this paper.

[6] In the TELRIC calculation, costs are based on the elements of the system needed to provide the service, including the total attributable costs of that element calculated as the incremental cost required to produce an extra unit of that service over the long run (where all elements of the system can be varied).

risk, as well as its systematic counterpart, affects the required rate of return when the rate base is subject to optimization. In fact, greater unsystematic risk compounds the impact of the systematic component of demand risk. For example, the firm's ORC falls if demand falls, since falling demand means that even more units of capacity are under-utilized. In contrast, rising demand only raises replacement cost to the point where all existing assets are fully utilized; larger increases have no further impact. This asymmetry means that increased unsystematic demand risk, which (by definition) has no effect on the covariance of demand with market returns, increases the covariance between the firm's ORC and market returns. This, in turn, raises the systematic risk of the firm's cash flows. Using simulations, we show that the effect on the firm's allowed rate of return is economically significant for reasonable parameter values.

Rate of return regulation is most plausible where entry is prevented and technologies are changing slowly, whereas incentive regulation more readily allows price regulation to co-exist with some entry and decentralized decisions about investment. Recently, incentive regulation has been preferred over rate of return regulation because, to an extent that depends on the particulars of the regime, it removes the link between prices and firm-specific costs and profits, thereby providing incentives for firms to behave efficiently.[7] The widespread adoption of incentive regulation renders it extremely important that its implementation does actually promote efficiency over other forms of regulation. The effect on investment is particularly important through its effect on dynamic efficiency.[8] Our model addresses the dynamic efficiency over time issue discussed by Littlechild (2003, pp. 304–306). We show that if the firm is forced to supply and is expected to break-even on new investment, then the rate base should equal ODV.[9]

Drawing on the analysis of TELRIC of Mandy and Sharkey (2003), Littlechild (2003) argues that the use of ORC shifts the risk of forecast errors to the regulated firm and raises its cost of capital because it is impossible to predict with accuracy the future path of cost, technology, and demand.[10] He goes on to say that acceptance of this risk by the firm may improve the prospects of competition because prices are much more stable where the firm takes on the regulatory-price setting risk. Hausmann (1997) uses the option to invest to argue that the use of TELRIC to price elements of a network underprices the economic cost of the services provided and will adversely affect investment. Jorde, Sidak, and Teece (2000) argue that the common practice of using TELRIC in pricing elements of telecommunications networks that are unbundled by mandate raises the cost of equity of firms that own the networks, and consequently reduces investment in these networks.

We set up our model in Section 2 and outline the various regulatory possibilities we consider

[7] The distinction between rate of return and incentive regulation and the critical issues relating to incentive regulation are pointed out by Newbery (1999, Chapter 2) and Laffont and Tirole (1993, Chapters 1 and 2).

[8] Baumol (2002) argues that the dynamic efficiency of investment is the dominant factor affecting welfare.

[9] Evans and Guthrie (2003) show that irreversibility and the requirement to supply imply that incentive regulation is efficient relative to rate of return regulation.

[10] Mandy and Sharkey (2003) explain the effect on cost recovery of the regulated firm when prices are set on the basis of TELRIC at shorter intervals than asset lives in a world of certainty.

in the following section. The analysis of Marshall et al. (1981) is adapted for irreversibility and applied to the different regulatory regimes in Section 4. We examine the dependence of demand risk allocation on regulatory policy in Section 5, where we present some numerical measures of the implications of our analysis. We conclude in Section 6.

2 The model

We consider a firm which faces uncertain future demand and capital prices.[11] In year t the firm requires an asset with capacity X_t to meet the needs of all its potential customers. Each new unit of capacity built by the firm in year t costs P_t. Investment in capacity is irreversible and the firm's assets are infinitely-lived. The firm has total revenue and operating cost in year t of R_t and C_t respectively.[12]

The firm is regulated in two ways. Firstly, the firm is required to meet the demand of all customers who wish to trade with the firm; that is, its capacity S_t must satisfy $S_t \geq X_t$ in each year t. We call this a universal service obligation.[13] Secondly, the regulator restricts the firm's revenue. The precise form of this restriction is not specified, but we assume that in each year t the regulator chooses the set of parameters A_t that influence the distribution of R_{t+1} and, possibly, X_{t+1}.[14] The form of regulation determines the nature of the parameters. For example, if a price cap is used, the parameters will comprise a starting output price cap and an adjustment to the inflation rate used to determine the future path of the cap (that is, amounting to a particular choice of 'X' in RPI-X regulation).

We assume that the firm has no flexibility in the timing of its investment, so it does not invest in capacity which is not needed to meet demand. In practice, the uncertainty surrounding future capital prices means that there may be instances in which the firm would choose to build excess capacity, such as when capital prices are expected to rise in the future. However, we eliminate such flexibility for several reasons. Firstly, this assumption means that excess capacity really

[11] Various factors can influence these two variables, but an important one for the industries we consider is technological change — technological advances can reduce demand (for example, increased use of mobile phones reducing demand for existing public switched telephone networks (PSTN)) or increase it (for example, greater demand for existing PSTN networks due to the advent of technologies such as ADSL), and can lead to dramatically lower capital prices.

[12] Since we are keeping our model abstract in order to illustrate the importance of irreversibility, and not the particular industry, for regulation, the precise interpretation of capacity will vary with the industry considered. For a telecommunications firm, capacity might represent the number of connections to the network; for an electricity distribution network, capacity might reflect the peak load carried over the network.

[13] Universal service obligations are usually motivated by income distributional concerns, although in telecommunications they can be motivated by network externalities whereby adding new customers to the network raises the welfare of others.

[14] The firm may have two or more distinct activities which are regulated. In this case, A_t might be a vector of regulated prices, and the distribution of R_{t+1} (which is the total revenue from all sources) will potentially depend on all these prices.

is "excess" — that is, if a hypothetical firm invested in assets to replace the regulated firm, it would not build the excess capacity. Secondly, there has been vigorous debate about the effect of investment flexibility (and other real options) on the revenue which regulated firms should be allowed to collect, with some authors arguing that a 'real option' premium should be added to the weighted-average cost of capital when calculating such firms' costs of capital. In this paper we show that a premium is appropriate even when the regulated firm has no real options. Lastly, this assumption keeps the model tractable.

The combination of irreversibility, the universal service obligation imposed on the firm, and the consequent lack of investment timing flexibility has important implications for investment behavior. If the firm has capacity $S_t \geq X_t$ in year t, it will only have to invest in year $t+1$ if $X_{t+1} > S_t$, in which case investment expenditure equals $I_{t+1} = P_{t+1}(X_{t+1} - S_t)$ and the assets' capacity will rise to $S_{t+1} = X_{t+1}$. If $X_{t+1} \leq S_t$, then the firm's investment is zero and the capacity of its assets remains at $S_{t+1} = S_t$. Thus, the capacity of its assets in year $t+i$ will equal

$$S_{t+i} = \max\{S_t, X_{t+1}, X_{t+2}, \ldots, X_{t+i}\}$$

and investment expenditure in that year will be

$$I_{t+i} = P_{t+i} \max\{X_{t+i} - \max\{S_t, X_{t+1}, X_{t+2}, \ldots, X_{t+i-1}\}, 0\}.$$

Investment in each future year is thus a non-increasing function of current capacity.

Although a firm building a network from scratch in year t would not build any excess capacity, it does not follow that excess capacity held by an existing firm has no value. In fact, if the regulated firm has excess capacity in year t, its future investment expenditure will sometimes be lower, and will never be greater, than that of a firm which starts business in year t. The latter firm incurs investment expenditure of

$$I^{(t)}_{t+i} = P_{t+i} \max\{X_{t+i} - \max\{X_t, X_{t+1}, \ldots, X_{t+i-1}\}, 0\}$$

in year $t+i$, so that the regulated firm's excess capacity reduces its investment expenditure in year $t+i$ by the amount

$$\begin{aligned} I^{(t)}_{t+i} - I_{t+i} &= P_{t+i} \max\{X_{t+i} - \max\{X_t, X_{t+1}, \ldots, X_{t+i-1}\}, 0\} \\ &\quad - P_{t+i} \max\{X_{t+i} - \max\{S_t, X_{t+1}, \ldots, X_{t+i-1}\}, 0\}, \end{aligned}$$

which is nonnegative. We will see that successful implementation of incentive regulation requires an understanding of the cost savings excess capacity can generate.

3 Rate bases

The market value of the regulated firm equals the market value of the net cash flows received by the firm's owners. These depend, in part, on the revenue which the regulator allows the firm to collect from its customers. Allowed revenue typically includes compensation for the firm's investment in physical capital, and this generally takes the form of the product of the "value" of the firm's assets and some regulated rate of return. Clearly this value cannot equal the market value of the regulated firm, as then a circularity results: the firm's market value depends on its allowed revenue, which depends on its market value. What is needed is some exogenous measure of the value of the firm's assets, known as the firm's rate base. In this paper we consider four different possibilities.[15]

Historical cost (HC). From year $t-1$ to year t, the firm invests I_t in new assets. Thus, the historical cost of the firm's assets evolves according to $HC_t = HC_{t-1} + I_t$.

Replacement cost (RC). If the firm was to replace its assets in year t, it would have to acquire S_t units of capacity at a price of P_t. Thus, replacement cost in year t is $RC_t = P_t S_t$.

Optimized replacement cost (ORC). If the firm was to replace its assets in year t with an optimal configuration, it would acquire $\min\{S_t, X_t\}$ units of capacity at a price of P_t, since only X_t units would be required if there was excess capacity, while all S_t units would have to be replaced if there was no excess capacity. Therefore, optimized replacement cost in year t is $ORC_t = P_t \min\{S_t, X_t\}$.

Optimized deprival value (ODV). This asset value measures the reduction in the value of the firm if it was deprived of its assets. The impact of such an event depends on whether or not the assets are currently being used to full capacity. If the firm is currently operating at full capacity, it would immediately rebuild S_t units of capacity if it lost its assets, and future investment expenditures would be unaffected by the loss. Thus, losing the assets would cost the firm $ODV_t = P_t S_t$. On the other hand, if the firm currently has excess capacity, then it would immediately rebuild just X_t units of capacity, costing $P_t X_t$, if it lost its assets. However, future investment expenditure would rise by $\hat{I}_{t+i}^{(t)} - I_{t+i} \geq 0$ in year $t+i$ for all $i \geq 1$. Thus, losing the assets would cost the firm

$$ODV_t = P_t X_t - M_t + U_t^{(t)},$$

where M_t denotes the value (measured in year t) of the firm's investment expenditure from year $t+1$ onwards, and $U_t^{(t)}$ denotes the value (also measured in year t) of all investment expenditure incurred from year $t+1$ onwards by a hypothetical efficient replacement firm that replaces the existing network in year t. This is the sum of the optimized replacement cost of the firm's assets and the value of the firm's excess capacity. Combining these two cases, we see that the optimized deprival value of the firm's assets is

$$ODV_t = ORC_t + (\text{value of excess capacity at } t). \tag{1}$$

[15] Our assumption that the firm's assets have infinite lives, made to keep the analysis as simple as possible, means that the issue of physical depreciation does not arise. The following definitions of rate bases need to be modified when assets have finite physical lives.

The total element long-run incremental cost (TELRIC) approach in telecommunications has been applied widely in Europe and the US. For the calculation of the TELRIC of a telecommunications service, costs are based on the elements of the system needed to provide the service, including the total attributable costs of that element calculated as the incremental cost required to produce an extra unit of the service over the long run (where all elements of the system can be varied). Thus, TELRIC fits our definition of ORC.

Regulators in Australia and New Zealand have considered a rate base which they term optimized deprival value, although their rate base has more in common with our ORC measure.[16] Current regulatory practice assigns each asset a value equal to its replacement cost, when the firm would replace the asset if it was deprived of the asset's use, and the so-called "economic value" of the asset otherwise. Economic value is defined to be the present value of the profit-maximizing revenue that could be generated from the asset. Since the asset would not be replaced if this present value is less than the asset's replacement cost, this rule leads to the following allowed value of an individual asset:

$$\text{contribution to rate base at } t = \min\{RC_t, \text{economic value at } t\}. \tag{2}$$

In our model, assets are either fully utilized or are not utilized at all, so that the economic value term in (2) is zero. Therefore, if $X_t \leq S_t$ only X_t units of capacity would be replaced (costing $P_t X_t$), and the remaining $S_t - X_t$ units of capacity have zero economic value, implying a rate base of $P_t X_t$. In contrast, if $X_t > S_t$ then all S_t units of capacity would be replaced, costing $P_t S_t$, which is the firm's rate base. That is, the optimized deprival value concept used in Australia and New Zealand corresponds to ORC.[17]

4 The approach of Marshall et al.

In our model the regulator sets prices and other regulatory parameters in order that the regulated firm can achieve some desired level of revenue. This level of revenue must be just sufficient to compensate the firm for the costs which it incurs, comprising operating costs and the cost of capital. The regulator must make three decisions: (1) What is the appropriate rate base? (2) What rate of return is the firm allowed to earn on this rate base? (3) What prices is the firm allowed to charge? Marshall et al. (1981) point out that the prices set by a regulator affect the risk borne by the firm, and therefore the rate of return which the firm should be allowed to

[16] Regulation of electricity transmission in New Zealand uses such a rate base (Ministry of Economic Development, 2000). Although considered favorably by regulators in Australia, ODV was rejected in favor of an ORC regime (Clarke, 1998; Johnstone, 2003, p. 3).

[17] The key omission from regulators' optimized deprival value calculations is the value of excess capacity. As a result, current practice does not measure the true effect on the value of a firm if it was deprived of its assets. Whereas current practice when calculating economic value is to use the present value of profit-maximizing revenue that the asset can generate, we add in the present value of the investment expenditures which will be avoided in the future because of the firm's ownership of the asset.

earn on its rate base, but do not consider irreversibility or the choice of rate base. They use the CAPM to determine the minimum revenue that investors require in order to participate. They argue that the regulator should set prices in such a way that the market value of the firm equals the cost of the firm's assets — investors would not be willing to participate if the market value is any lower. In this section we describe how Marshall et al.'s approach would be implemented for the types of firms we consider in this paper, firstly for an arbitrary choice of rate base, and then for the four rate bases described in Section 2.

The regulator sets prices and other regulatory parameters A_t such that the firm exactly breaks even when it first builds the asset — that is, the zero-profit condition essential for Ramsey pricing is met. It does so by setting A_t such that the market value of the firm is always equal to some exogenous rate base. Provided that the selected rate base equals zero whenever the firm's assets have zero capacity, the firm breaks even when it first builds the asset. The specific choice of rate base then determines how the value of the firm evolves in the future.

If the market value of the firm is F_{t+1} at the end of year $t+1$, then the value of investors' stake in the firm at the start of year $t+1$ is

$$V_{t+1} = R_{t+1} - C_{t+1} - I_{t+1} + F_{t+1}.$$

From the certainty equivalent form of the CAPM, the most that investors will be willing to pay in year t for their stake in the firm is

$$F_t(A_t) = \frac{E_t[V_{t+1}|A_t] - \lambda_t \mathrm{Cov}_t[V_{t+1}, r_{m,t}|A_t]}{1 + r_{f,t}},$$

where $\lambda_t = (E_t[r_{m,t}] - r_{f,t})/\mathrm{Var}_t[r_{m,t}]$ is the market price of risk, $r_{f,t}$ is the risk-free interest rate over the year $[t, t+1]$, $r_{m,t}$ is the (risky) rate of return on the market portfolio over the same period, and the regulator's choice of A_t affects the indicated expectations. The value of the firm will equal the desired rate base B_t if the regulator sets regulatory parameters such that

$$B_t = \frac{E_t[R_{t+1} - C_{t+1} - I_{t+1} + B_{t+1}|A_t] - \lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} - I_{t+1} + B_{t+1}, r_{m,t}|A_t]}{1 + r_{f,t}}. \tag{3}$$

Solving for $E_t[R_{t+1}|A_t]$ shows that expected revenue must equal

$$\begin{aligned} E_t[R_{t+1}|A_t] &= (1 + r_{f,t})B_t + \lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} - I_{t+1} + B_{t+1}, r_{m,t}|A_t] \\ &\quad + E_t[C_{t+1}|A_t] + E_t[I_{t+1}|A_t] - E_t[B_{t+1}|A_t]. \end{aligned} \tag{4}$$

By choosing A_t, the regulator selects the distribution from which revenue R_{t+1} will be drawn. We will not consider the price-setting decision explicitly, but rather assume that the resulting distribution of revenue is consistent with (4).[18]

[18] Multiple prices may arise because the regulated firm has several lines of business (not all of which need be regulated) or because the regulator has subdivided markets and imposed different prices for the same good. Equation (4) gives a single condition relating the different prices, not one equation for each line of business. It is the only restriction on the regulator's choices that is imposed by the requirement that the value of the firm equal the regulator's chosen rate base.

It is convenient to rewrite (4) in the following form:[19]

Proposition 1 *If the value of the regulated firm is to equal its rate base B_t, then its expected revenue must satisfy*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t B_t + (B_t - E_t[B_{t+1} - I_{t+1}]), \tag{5}$$

where the allowed rate of return equals

$$r_t = r_{f,t} + \lambda_t Cov_t \left[\frac{R_{t+1} - C_{t+1} - I_{t+1} + B_{t+1}}{B_t}, r_{m,t} \right]. \tag{6}$$

The expected revenue in (5) can be naturally decomposed into three terms. The first, $E_t[C_{t+1}]$, is the firm's expected operating cost at the end of the year. The second term, $r_t B_t$, equals the amount which the firm's investors are allowed to earn on the rate base. From (6), the allowed rate of return is the sum of the risk free interest rate and a risk premium which compensates investors for the systematic risk of shocks to the future value of their stake in the firm. The third term, $B_t - E_t[B_{t+1} - I_{t+1}]$, can be interpreted as the expected reduction in value of the investors' asset, or expected economic depreciation: provided investment raises the rate base dollar-for-dollar, then the market value of their year t investment changes from B_t in year t to $B'_{t+1} = B_{t+1} - I_{t+1}$ in year $t + 1$.[20]

The value in year $t + 1$ of investors' stake in the firm is

$$\begin{aligned} V_{t+1} &= R_{t+1} - C_{t+1} - I_{t+1} + B_{t+1} \\ &= (1 + r_t)B_t + ((R_{t+1} - C_{t+1}) - (E_t[R_{t+1} - C_{t+1}])) + (B'_{t+1} - E_t[B'_{t+1}]). \end{aligned}$$

When viewed from year t, investors are thus exposed to shocks to net revenue, $R_{t+1} - C_{t+1}$, and shocks to the value of the current rate base, B'_{t+1}. These shocks affect the firm in fundamentally different ways. Net revenue shocks that occur in the middle of the regulatory cycle (for example, operating costs being higher than expected) are only costly to the firm until prices are reset. In contrast, shocks to the rate base only directly affect the firm once prices are reset, when they are permanent.

We now consider the four rate bases described in Section 3. In all cases, investors are exposed to the risk of net revenue shocks. However, the regimes expose investors to different risks of rate base shocks.

4.1 Historical cost

Under traditional rate of return regulation, the rate base rises each year by the amount of investment made by the regulated firm. That is, if the rate base in year t equals B_t, then the

[19] In order to simplify the notation as much as possible, from now on we suppress the notation indicating that expectations are conditional on A_t.

[20] Mandy and Sharkey (2003) analyze revenue requirements that allow for economic depreciation, but, because they assume an exogenous cost of capital, they do not consider the effect on risk of the choice of rate base.

rate base in year $t+1$ equals $B_{t+1} = B_t + I_{t+1}$, the historical cost of the firm's assets. Since this means that $B'_{t+1} = B_t$, investors are only exposed to the risk of shocks to net revenue. Substituting the expression for B_{t+1} into (5) and (6) proves

Proposition 2 *When the historical cost of the firm's assets is adopted as the rate base, the firm's expected revenue must equal*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t B_t,$$

where the allowed rate of return is

$$r_t = r_{f,t} + \lambda_t \mathrm{Cov}_t \left[\frac{R_{t+1} - C_{t+1}}{B_t}, r_{m,t} \right].$$

The firm should not receive any compensation for expected economic depreciation because no such depreciation can occur. The allowed rate of return reflects the fact that the only risk the firm is exposed to is that shocks to net revenue might occur after the regulator sets prices (and before it resets them in year $t+1$).

4.2 Replacement cost

When the replacement cost of the firm's assets is chosen as the rate base, $B_t = P_t S_t$, $B_{t+1} = P_{t+1} S_{t+1}$, and $I_{t+1} = P_{t+1}(S_{t+1} - S_t)$. Therefore $B'_{t+1} = P_{t+1} S_t$, and the only source of rate base risk is shocks to the capital price. Substituting the expressions for B_{t+1} and I_{t+1} into (5) and (6) proves

Proposition 3 *When the replacement cost of the firm's assets is adopted as the rate base, the firm's expected revenue must equal*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t B_t + (P_t - E_t[P_{t+1}]) S_t, \tag{7}$$

where the allowed rate of return is

$$r_t = r_{f,t} + \lambda_t \mathrm{Cov}_t \left[\frac{R_{t+1} - C_{t+1}}{B_t} + \frac{P_{t+1}}{P_t}, r_{m,t} \right]. \tag{8}$$

As in the case of an historical cost rate base, investors should be compensated for net revenue risk. Now, however, they should also be compensated for fluctuations in the capital price: compensation for anticipated changes comes via the economic depreciation term, $(P_t - E_t[P_{t+1}])S_t$, in (7); compensation for unanticipated changes comes through the additional risk premium,

$$\lambda_t \mathrm{Cov}_t \left[\frac{P_{t+1}}{P_t}, r_{m,t} \right],$$

in (8). Apart from ex ante compensation in the form of the additional risk premium, investors bear the full cost of unanticipated changes in capital prices.

4.3 Optimized replacement cost

The regulated firm is exposed to an additional source of rate base risk when the optimized replacement cost of its assets is chosen as the rate base. To see why, note that investment is necessary in year $t+1$ if and only if $X_{t+1} > S_t$, in which case investment expenditure will equal $I_{t+1} = P_{t+1}(X_{t+1} - S_t)$. Since $B_{t+1} = P_{t+1}X_{t+1}$ for all values of X_{t+1}, the future value of the current rate base is

$$B'_{t+1} = B_{t+1} - I_{t+1} = P_{t+1} \min\{S_t, X_{t+1}\}.$$

Thus, rate base risk arises through shocks to both capital prices and demand, and this is reflected in the expected revenue which the firm should be allowed to collect. Substituting the expressions for B_{t+1} and I_{t+1} into (5) and (6) proves[21]

Proposition 4 *When the optimized replacement cost of the firm's assets is adopted as the rate base, the firm's expected revenue must equal*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t B_t + \big(P_t X_t - E_t[P_{t+1} \min\{S_t, X_{t+1}\}]\big),$$

where the allowed rate of return is

$$r_t = r_{f,t} + \lambda_t \mathrm{Cov}_t \left[\frac{R_{t+1} - C_{t+1}}{B_t} + \frac{P_{t+1}}{P_t} \min\left\{ \frac{S_t}{X_t}, \frac{X_{t+1}}{X_t} \right\}, r_{m,t} \right].$$

As was the case with a replacement cost rate base, the firm is exposed to the risk of capital price shocks, but now it also faces demand risk. Demand risk is asymmetric, since increases in demand beyond the assets' current capacity have no additional impact on the firm's rate base, while there is unlimited downside risk from negative demand shocks. This asymmetry means that expected economic depreciation depends on the variability of future demand, as well as on its mean. The firm receives full compensation for anticipated changes in demand, through the allowance for expected economic depreciation, while the only compensation for unanticipated changes is received ex ante, via the risk premium.

When replacement cost is used as the rate base, the firm's allowed revenue is decreasing in the excess capacity of the firm's assets — the firm is punished for having excess capacity. This is inconsistent with modern incentive regulation, which requires that the firm's allowed revenues depend only on the cost structure of an efficient alternative provider. Thus incentive regulation is unable to drive the market value of the regulated firm to the replacement cost of an efficiently-configured firm.

[21] When investment is reversible we obtain the same expressions for expected revenue and allowed rate of return as in Section 4.2, because the firm can simply sell (or redeploy) any excess capacity following a negative demand shock — the firm is exposed to the risk of fluctuations in the capital price, but not in the quantity of capital required. Since the firm will never carry excess capacity, $S_t = X_t$ at all times. If demand rises from year t to year $t+1$, the firm will have to invest $P_{t+1}(X_{t+1} - X_t)$ in new capacity; if it falls from year t to year $t+1$, the firm will raise $P_{t+1}(X_t - X_{t+1})$ from selling excess capacity. Overall, investment is $I_{t+1} = P_{t+1}(X_{t+1} - X_t)$. Substituting these values into (5) and (6), together with the new rate base $B_t = P_t X_t$, results in (7) and (8).

4.4 Optimized deprival value

In the final possibility we consider, the regulated firm receives the level of revenue that would be required for a hypothetical firm (which can invest in assets configured to meet current and future demand) to replace the incumbent and just break even. This is true incentive regulation — the regulated firm's allowed revenue is determined by the cost structure of a hypothetical efficient replacement firm. If such a replacement firm invests in year t and undertakes all future investment required to meet the universal service obligation, then its net cash flow is $-P_t X_t$ in year t, and $R_{t+i} - C_{t+i} - \hat{I}^{(t)}_{t+i}$ in year $t+i$ for all $i \geq 1$. The regulator sets prices such that the market value of all cash flows from year $t+1$ onwards equals $P_t X_t$, so that the hypothetical replacement firm just breaks even. Relative to this firm, the regulated firm receives additional cash flows of $\hat{I}^{(t)}_{t+i} - I_{t+i}$ in year $t+i$ for all $i \geq 1$. Thus, the market value of the regulated firm exceeds the market value of the hypothetical replacement firm, $P_t X_t$, by an amount equal to the market value of the stream of incremental cash flows resulting from assets unused now but usable in the future. It follows that the market value of the regulated firm equals the sum of the ORC of its assets, $P_t X_t$, and the value of its excess capacity; from (1), this equals the ODV of its assets.

Substituting the firm's ODV as the rate base in (5) and (6) proves

Proposition 5 *When the optimized deprival value of the firm's assets is adopted as the rate base, the expected revenue must equal*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t ODV_t + (ODV_t + E_t[I_{t+1}] - E_t[ODV_{t+1}]),$$

where the allowed rate of return is

$$r_t = r_{f,t} + \lambda_t \mathrm{Cov}_t \left[\frac{R_{t+1} - C_{t+1} - I_{t+1} + ODV_{t+1}}{ODV_t}, r_{m,t} \right].$$

Since the firm's ODV depends on the actual capacity of its assets, and ODV appears in both equations in Proposition 5, it might seem that the firm's allowed revenue depends on this capacity. However, the discussion preceding Proposition 5 shows that this is not the case — the firm's expected revenue is independent of S_t, as is required for incentive regulation. In the appendix we prove

Corollary 1 *When the optimized deprival value of the firm's assets is adopted as the rate base, the expected revenue must equal*

$$E_t[R_{t+1}] = E_t[C_{t+1}] + r_t P_t X_t + \left(P_t X_t - E_t[P_{t+1} \min\{X_t, X_{t+1}\} + U^{(t+1)}_{t+1} - U^{(t)}_{t+1}] \right), \quad (9)$$

where the allowed rate of return is

$$r_t = r_{f,t} + \lambda_t Cov_t \left[\frac{R_{t+1} - C_{t+1}}{P_t X_t} + \frac{P_{t+1}}{P_t} \min\left\{1, \frac{X_{t+1}}{X_t}\right\} + \frac{U^{(t+1)}_{t+1} - U^{(t)}_{t+1}}{P_t X_t}, r_{m,t} \right] \quad (10)$$

and $U^{(s)}_{t+1}$ denotes the market value, measured in year $t+1$, of all investment expenditure from year $t+2$ onwards of a hypothetical firm that started business in year s.

Corollary 1 shows that incentive regulation can be achieved using ORC as the rate base, but not if we use a naive implementation of Marshall et al.'s approach. Instead, allowance must be made for the possibility that some current, efficiently-configured, capacity becomes temporarily under-utilized in the future. The value of this future excess capacity affects both expected economic depreciation and the allowed rate of return. This is important, since, as we discussed in Section 3, regulators do not include the value of excess capacity in ODV calculations.

We finish this section by illustrating how ODV is calculated when, as is common, regulated firms have two or more distinct activities utilizing their rate base assets, only some of which are regulated.[22] We assume, for simplicity, that the unregulated activities do not raise the firm's operating costs and ask what level of revenue would be required for a hypothetical firm to replace the incumbent and just break even. Suppose the unregulated activities generate revenue of R_t^u and $\hat{R}_t^u$ in year t for the regulated firm and its hypothetical replacement, respectively. Suppose that both firms are allowed to collect revenue of R_t^r from the regulated activity. If a replacement firm invests in year t and undertakes all future investment required to meet the universal service obligation, then its net cash flow is $-P_t X_t$ in year t, and $R_{t+i}^r + \hat{R}_{t+i}^u - C_{t+i} - \hat{I}_{t+i}^{(t)}$ in year $t+i$ for all $i \geq 1$. The regulator sets prices such that the market value of all cash flows from year $t+1$ onwards equals $P_t X_t$, so that the hypothetical replacement firm just breaks even. Relative to this firm, the regulated firm receives cash flows of $R_{t+i}^u - \hat{R}_{t+i}^u + \hat{I}_{t+i}^{(t)} - I_{t+i}$ in year $t+i$ for all $i \geq 1$. The market value of the regulated firm exceeds the market value of the hypothetical replacement firm, $P_t X_t$, by an amount equal to the market value of the stream of incremental cash flows. It follows that the market value of the regulated firm equals the sum of $P_t X_t$, the ORC of its assets, and the value of its excess capacity — where the value of excess capacity is now the sum of the value of the future investment expenditure avoided and the value of any additional profits which the firm can earn from unregulated activities because it owns this excess capacity. With this extended definition of ODV, the market value of the regulated firm equals the ODV of its assets.[23]

5 Demand risk

5.1 Allocating demand risk between the firm and consumers

From Propositions 2 and 3, when the regulator chooses historical or replacement cost as the firm's rate base the allowed expected revenue is independent of demand — a drop in demand must be met with a higher regulated price. However, demand might fall to such a low level

[22] For example, if portions of a firm's network are judged to be 'bottleneck facilities', only the wholesale price that the firm may charge its competitors for access to the bottleneck facility may be regulated, while its retail prices are unregulated.

[23] Note that the regulated firm effectively gets to keep only the additional profits from unregulated activities that derive from its excess capacity; any profits it would receive from unregulated activities on an efficiently-configured network reduce the revenue allowed from the regulated activity.

that this required level of revenue cannot be collected by the regulated firm; that is, it may not be possible to solve (4) for the regulatory parameters A_t. In contrast, Propositions 4 and 5 show that as soon as the firm's asset base is subject to optimization by the regulator, the firm is exposed to the risk of demand shocks — with both ORC and ODV rate bases the allowed expected revenue will also fall if demand falls. Thus, ORC and ODV are more credible rate bases when future demand is potentially subject to large negative shocks from, for example, competitive entry or technological advances.

5.2 Allocating demand risk across consumers

The universal service obligation, investment irreversibility, and customers' options to abandon the network create the possibility of wealth transfers between customers. Suppose, for example, that an additional customer joins the network in year t, but leaves it in year $t+1$, and that historical cost is used as the rate base. If investment is required to allow this customer to connect to the network in year t, then the rate base the following year is raised by P_t. Since the firm is guaranteed to recover all of its investment costs, the firm is unaffected by the customer's departure. However, the customers still connected to the network in year $t+1$ will have to pay more — in effect, those customers who remain pay for (at least part of) the abandonment options of those who disconnect. Similar transfers occur with replacement cost and ORC rate bases, since in both cases allowed revenue in year $t+1$ depends on the actual level of capacity in year t.

However, when ODV is used as the rate base, the allowed revenue in year $t+1$ and beyond is independent of the actual level of demand in year t. Thus, the revenue collected from those customers who remain is unaffected by the presence (or absence) of customers who joined the network in the past — with an ODV rate base, customers pay for their option to abandon the network ex ante. It follows that, with ODV as the rate base, a customer does not impose any costs on other parties when it joins the network: (i) although the firm is exposed to the risk of early departure by the customer, it is compensated ex ante for bearing this risk; (ii) other customers are not affected if the customer subsequently leaves the network.

5.3 Compensating the firm for its exposure to demand risk

In this section we investigate the determinants of the allowed rate of return of a regulated firm with an ODV rate base. Because the precise level of the firm's ODV depends on the distribution of all future levels of demand, we cannot derive simple expressions involving (X_t, X_{t+1}) for the firm's expected revenue and allowed rate of return under ODV-regulation. However, simple functional forms are possible in a special case of our model in which required capacity may change from year t to year $t+1$ but remains constant at its new level after year $t+1$. This means that no investment is required from year $t+2$ onwards, implying that $U_{t+1}^{(t+1)} = U_{t+1}^{(t)}$. Substituting this into (10) shows that the regulated firm must be allowed to earn the rate of

Table 1: The effect of demand risk on the allowed rate of return

β	σ_ϵ			
	0.0	0.1	0.2	0.3
0.00	0.0000	0.0000	0.0000	0.0000
0.25	0.0055	0.0079	0.0089	0.0093
0.50	0.0110	0.0134	0.0158	0.0170
0.75	0.0165	0.0186	0.0216	0.0237

Notes. The table reports the risk premium resulting from the risk of demand shocks. The return on the market portfolio is normally distributed with mean 0.12 and standard deviation 0.2, and the growth in demand equals $X_{t+1}/X_t - 1 = \beta r_{m,t} + \epsilon_t$, where ϵ_t is normally distributed with mean zero and standard deviation σ_ϵ. Demand shocks thus have systematic risk of β. The riskless interest rate is 0.04. The covariances required to calculate the risk premia are each estimated using simulations with 100,000 draws.

return

$$r_t = r_{f,t} + \lambda_t \mathrm{Cov}_t \left[\frac{R_{t+1} - C_{t+1}}{P_t X_t} + \frac{P_{t+1}}{P_t} \min \left\{ 1, \frac{X_{t+1}}{X_t} \right\}, r_{m,t} \right]$$

on the ORC of its assets.

We use numerical simulations to show the significant impact of demand risk (both systematic and unsystematic) on the allowed rate of return, simplifying our analysis by assuming that the capital price is constant. For a range of parameter values, we estimate the component of the risk premium due to demand shocks,

$$\lambda_t \mathrm{Cov}_t \left[\min \left\{ 1, \frac{X_{t+1}}{X_t} \right\}, r_{m,t} \right].$$

The return on the market portfolio is normally distributed with mean 0.12 and standard deviation 0.2, and the growth in demand equals

$$\frac{X_{t+1}}{X_t} - 1 = \beta r_{m,t} + \epsilon_t, \tag{11}$$

where ϵ_t is normally distributed with mean zero and standard deviation σ_ϵ. Demand shocks thus have systematic risk of β. The riskless interest rate is 0.04. The covariances required to calculate the risk premia are each estimated using simulations with 100,000 draws, and the results of this exercise are reported in Table 1. For example, if demand shocks have systematic risk of $\beta = 0.5$ and the unsystematic component of demand shocks has a standard deviation of $\sigma_\epsilon = 0.1$, the appropriate risk premium is 1.34 percent.

The risk premium for demand risk, which can be economically significant, is an increasing function of systematic demand risk. Further, if demand has positive systematic risk, any additional unsystematic demand risk adds to the risk premium because of the nonlinearity introduced by the irreversibility of investment and the requirement to supply. To illustrate this, consider (11). It includes the elements of demand correlated with $r_{m,t}$ (that is,

$E[X_{t+1}/X_t - 1|r_{m,t}] = \beta r_{m,t})$ and the uncorrelated elements represented by ϵ_t. The otherwise unsystematic elements, ϵ_t, enter systematic risk under regulation because they affect the truncation point; using equation (10),

$$\mathrm{Cov}_t\left[\min\left\{1, \frac{X_{t+1}}{X_t}\right\}, r_{m,t}\right] = \mathrm{Cov}_t[\beta r_{m,t} + \epsilon_t, r_{m,t}|\beta r_{m,t} + \epsilon_t < 0].$$

Provided $\beta \neq 0$, this covariance is clearly a function of the distribution of ϵ_t.

In short, when installed assets are sunk and the firm is required to supply, regulatory regimes that use ODV (or ORC) to set prices impart additional volatility to the firm's future value. This raises the revenue that the firm must be allowed to earn if it is to break even on new investment. The higher revenue is to cover the expected cost of assets under-utilized in the future, as well as the higher returns needed to compensate for the increased risk due to capital price and demand uncertainty. In addition, the truncation of revenues implied by irreversibility means that both systematic and unsystematic demand risk affect the allowed rate of return.

6 Concluding remarks

We have identified various ways in which regulators can ensure that regulated firms just break even whenever they are forced to make irreversible investments in infrastructure. This condition, which extends Ramsey pricing to a dynamic setting, is necessary if provision of infrastructure is to be sustainable over time. The different approaches vary according to the rate base set by the regulator. One rate base, optimized deprival value (ODV), leads to allowed revenue which is forward-looking — that is, which does not depend on past investment decisions. In contrast, we show that if more traditional rate bases, such as historical cost and optimized replacement cost (ORC), are implemented in a way that treats sunk costs sustainably (that is, in such a way that the firm breaks even when it is forced to invest) then allowed revenue is backward-looking — that is, it depends on past demand (via its dependency on past investment decisions).

Our work shows that if incentive regulation is to be sustainable over time then the rate base should be ODV, which exceeds ORC by an amount equal to the present value of expected cost savings resulting from the firm's 'excess' capacity. Further, if firms are to break even on new investments then the allowed rate of return should be sufficient to compensate for the extra risks implied by incentive regulation. Technology shocks are especially important since they can affect both capital prices and demand. For both ODV and ORC rate bases, the firm must be compensated for anticipated changes in technology through an allowance for expected economic depreciation; it must be compensated, ex ante, for unanticipated changes through the allowed rate of return. Systematic risk may be increased by the introduction of specific unsystematic risks because of the truncation resulting from irreversibility and the requirement to supply. Not using this rate base specification or allowing for these risks will incorrectly compensate the firm for its investment and thereby adversely affect investment and dynamic efficiency.[24]

[24] We thus disagree with those who would set prices based on TELRIC calculations at intervals on the grounds

There are two sets of measurement issues suggested by our analysis. The first is that it is not sufficient to draw comparator firms from the same industry where benchmark comparisons are made in the process of setting regulatory parameters — they must also be subject to the same regulatory environment. For example, European and US firms' rates of return under regulation, even if drawn from the same industry, will likely differ as a result of different regulatory regimes. The second measurement issue relates to assessing the profitability of regulated industries. Those industries that are subject to rate of return regulation can be expected to have lower ex ante and ex post rates of return than those subject to incentive regulation. Under incentive regulation, the risks of demand and capital price fluctuations are borne by the firm rather than consumers; the reverse is true for rate of return regulation. This difference between the regimes is exacerbated where regulation seeks to enhance competition and thereby introduces further risk to the incumbent's business.

Although it is a policy issue for the firm and regulators, and a topic deserving further research, we have not investigated the decision to invest in advance of demand. Where there are economies to be achieved by investing in asset configurations for which capacity is not expected to be attained until some future date, such investment may be economically desirable. We conjecture that use of the ODV rate base would allow the firm to just break even if it could include in its ODV assets that are installed in anticipation of demand. These assets would be assigned a value equal to the potential cost savings that they implied, leaving the firm the cost saving as inducement for efficient investment prior to demand; assets for which there is no prospective use would have no future cost savings, and hence contribute nothing to ODV. The allowance of future cost savings associated with the calculated ODV is critical to specifying the rate base that allows the expected recovery of investment, and it has a significant speculative component that the regulator will generally want to assess. In this respect, and in the verification of ODV more generally, these sorts of incentive regimes may require detailed monitoring and regulator-decision-making that approach those of rate of return regulation.

References

Asplund, Marcus (2000). "What fraction of a capital investment is sunk costs?" *Journal of Industrial Economics* 48(3), 287–304.

Baumol, William J. (2002). *The Free-market Innovation Machine: Analyzing the Growth Miracle of Capitalism*, Princeton, New Jersey: Princeton University Press.

Clarke, Frank L. (1998). "Deprival value and optimized deprival value in Australasian public sector accounting: Unwarranted drift and contestable serviceability," *Abacus* 34(1), 8–17.

of dynamic efficiency.

Cowan, Simon (2003). "Optimal risk allocation for regulated monopolies and consumers," *Journal of Public Economics* 88, 285–303.

Economides, Nicholas (1999). "Real options and the costs of the local telecommunications network." In: Alleman, James, and Eli Noam (eds.), *The New Investment Theory of Real Options and its Implications for the Cost Models in Telecommunications*, Kluwer Academic Publishers, Boston, Massachusetts.

Evans, Lewis T., and Graeme A. Guthrie (2003). "Stranding is inevitable: Implications for optimal regulatory design," Working paper, NZ Institute for the Study of Competition and Regulation.

Hausmann, Jerry A. (1997). "Valuing the effect of regulation on new services in telecommunications,"*Brookings Papers on Economic Activity: Microeconomics* 1, Brookings.

Hausman, Jerry A. (1999). "The effect of sunk costs in telecommunications regulation," In: Alleman, James, and Eli Noam (eds). *The New Investment Theory of Real Options and its Implications for Telecommunications Economics*, Kluwer Academic Publishers, Boston, Massachusetts.

Hausman, Jerry, and Stewart Myers (2002). "Regulating the United States railroads: The effects of sunk costs and asymmetric risk," *Journal of Regulatory Economics* 22(3), 287–310.

Johnstone, David J. (2003). "Replacement cost asset valuation and regulation of energy infrastructure tariffs," *Abacus* 39(1), 1–41.

Jorde, Thomas M., J. Gregory Sidak, and David J. Teece (2000). "Innovation, investment, and unbundling," *Yale Journal on Regulation* 17(1), 1–37.

Laffont, Jean-Jacques, and Jean Tirole (1993). *A Theory of Incentives and Procurement and Regulation*, MIT Press, Cambridge, Massachusetts.

Littlechild, Stephen (2003). "Reflections on incentive regulation," *Review of Network Economics* 2(4), 289–315.

Mandy, David M., and William W. Sharkey (2003). "Dynamic pricing and investment from static proxy models," *Review of Network Economics* 2(4), 403–439.

Marshall, William J., Jess B. Yawitz, and Edward Greenberg (1981). "Optimal regulation under uncertainty," *Journal of Finance* 36(4), 909–921.

Ministry of Economic Development (2000). *Handbook for Optimised Deprival Valuation of System Fixed Assets of Electricity Lines Businesses* (4th edition), Ministry of Economic Development, Wellington, New Zealand.

Newbery, David M. (1999). *Privatisation, Restructuring, and Regulation of Network Utilities*, MIT Press, Cambridge, Massachusetts.

Ramey, Valerie A., and Matthew D. Shapiro (2001). "Displaced capital: A study of aerospace plant closings," *Journal of Political Economy* 109(5), 958–992.

Spulber, Daniel F. (1989). *Regulation and Markets* MIT Press, Cambridge, Massachusetts.

Appendix

Proof of Corollary 1

Substituting the firm's ODV in year t, $P_t X_t + U_t^{(t)} - M_t$, into (3) shows that

$$\begin{aligned} P_t X_t + U_t^{(t)} - M_t &= \frac{E_t[R_{t+1} - C_{t+1} - I_{t+1} + P_{t+1}X_{t+1} + U_{t+1}^{(t+1)} - M_{t+1}]}{1 + r_{f,t}} \\ &\quad - \frac{\lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} - I_{t+1} + P_{t+1}X_{t+1} + U_{t+1}^{(t+1)} - M_{t+1}, r_{m,t}]}{1 + r_{f,t}}. \end{aligned} \tag{A-1}$$

Because M_t denotes the value in year t of all of the firm's future investment expenditure,

$$M_t = \frac{E_t[I_{t+1} + M_{t+1}] - \lambda_t \mathrm{Cov}_t[I_{t+1} + M_{t+1}, r_{m,t}]}{1 + r_{f,t}}.$$

Using this to eliminate M_t from (A-1) shows that

$$\begin{aligned} P_t X_t + U_t^{(t)} &= \frac{E_t[R_{t+1} - C_{t+1} + P_{t+1}X_{t+1} + U_{t+1}^{(t+1)}]}{1 + r_{f,t}} \\ &\quad - \frac{\lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} + P_{t+1}X_{t+1} + U_{t+1}^{(t+1)}, r_{m,t}]}{1 + r_{f,t}}. \end{aligned} \tag{A-2}$$

The definitions of $U_t^{(t)}$ and $U_{t+1}^{(t)}$ imply that

$$U_t^{(t)} = \frac{E_t[\hat{I}_{t+1}^{(t)} + U_{t+1}^{(t)}] - \lambda_t \mathrm{Cov}_t[\hat{I}_{t+1}^{(t)} + U_{t+1}^{(t)}, r_{m,t}]}{1 + r_{f,t}}.$$

Using this to eliminate $U_t^{(t)}$ from (A-2) shows that

$$\begin{aligned} P_t X_t &= \frac{E_t[R_{t+1} - C_{t+1} + P_{t+1}X_{t+1} - \hat{I}_{t+1}^{(t)} - U_{t+1}^{(t)} + U_{t+1}^{(t+1)}]}{1 + r_{f,t}} \\ &\quad - \frac{\lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} + P_{t+1}X_{t+1} - \hat{I}_{t+1}^{(t)} - U_{t+1}^{(t)} + U_{t+1}^{(t+1)}, r_{m,t}]}{1 + r_{f,t}} \\ &= \frac{E_t[R_{t+1} - C_{t+1} + P_{t+1}\min\{X_t, X_{t+1}) + U_{t+1}^{(t+1)} - U_{t+1}^{(t)}]}{1 + r_{f,t}} \\ &\quad - \frac{\lambda_t \mathrm{Cov}_t[R_{t+1} - C_{t+1} + P_{t+1}\min\{X_t, X_{t+1}) + U_{t+1}^{(t+1)} - U_{t+1}^{(t)}, r_{m,t}]}{1 + r_{f,t}}, \end{aligned}$$

where we have used the fact that $P_{t+1}X_{t+1} - \hat{I}_{t+1}^{(t)} = P_{t+1}\min\{X_t, X_{t+1})$ in the final step. Solving this equation for $E_t[R_{t+1}]$ results in (9) and (10).