Incentive Regulation of Prices when Costs are Sunk*

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

May 20, 2004

*We thank participants at an NZISCR seminar and the Advanced Workshop in Regulation and Competition's 16^{th} Annual Western Conference for helpful comments.

†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

Incentive Regulation of Prices when Costs are Sunk

Abstract

We present a model featuring irreversible investment, uncertain future demand and capital prices, and a regulator who sets the firm's output price at discrete intervals. Using this model, we derive a closed-form solution for the firm's output price which ensures that, whenever the regulator resets the price, the present value of the firm's future net revenue stream equals the present value of the investment expenditure incurred by a hypothetical efficient firm which replaced the regulated firm. We calculate the rate of return which shareholders should receive to compensate them for the exposure to demand risk and capital price risk induced by modern incentive regulation. In contrast to rate of return regulation, we find that resetting the regulated price more frequently increases the risk faced by the firm's owners, and that this is reflected in a higher output price and a higher weighted-average cost of capital. We show that the market value of the regulated firm will generally exceed the replacement cost of its existing assets by an amount that we interpret as the value of the firm's excess capacity. The higher valuation is required in order for the firm to prospectively manage fixed costs that are implied by irreversibility. We suggest it is indicative of the efficient treatment of investment in advance. This contrasts with much of the existing literature which argues that the market value of a regulated firm should equal the cost of its existing assets.

JEL Classification code: G31, L5

Keywords: Incentive regulation, uncertainty, sunk costs

Incentive Regulation of Prices when Costs are Sunk

1 Introduction

Much investment in infrastructure is irreversible, and the demand for infrastructure assets fluctuates over time. These two features combine to create the risk that irreversible investments will become unproductive before the end of their physical lives. Under traditional rate of return regulation, the regulated firm is allowed to adjust its prices so that it can recover the cost of its investments, even those assets which are unproductive, meaning that consumers bear the risk of adverse demand shocks.[1] The deregulation which took place in many countries in the 1980s and 1990s saw regulators remove unproductive assets from regulated firms' rate bases, effectively preventing the firms from recovering the cost of their investment in these assets, and shifting much of the demand risk onto firms' shareholders. Modern incentive regulation allows regulated firms to collect just enough revenue to cover the costs faced by a hypothetical 'efficient' replacement firm, thereby exposing regulated firms to the additional risk that capital prices, and therefore their own allowed revenue, will fall in the future. For example, the FCC's starting point in its TELRIC calculation is the cost structure of an efficient cost-minimizing firm with an optimally-configured network built with the current technology (Weisman, 2002).

We present a model featuring irreversible investment, uncertain future demand and capital prices, and a regulator who sets the firm's output price at discrete intervals. Using this model, we determine the reward, in the form of the rate of return allowed by the regulator, which shareholders should receive for the exposure to demand risk and capital price risk induced by incentive regulation. Specifically, we derive a closed-form solution for the firm's output price which ensures that, whenever the regulator resets the price, the present value of the firm's future net revenue stream equals the present value of the investment expenditure incurred by a hypothetical efficient firm which replaced the regulated firm. The regulated price depends on the systematic risk of demand and capital price shocks, but other factors, including the extent of unsystematic risk, are also important. In contrast to rate of return regulation, we find that resetting the regulated price more frequently increases the risk faced by the firm's owners, and that this is reflected in a higher regulated price.

In our model, the regulated firm's revenues are determined by the cost structure of the hypothetical lowest cost provider. We show that this implies that the market value of the regulated firm will generally exceed the replacement cost of its existing assets. The difference, which we interpret as the value of the firm's excess capacity, derives from the future cost savings which arise from assets which are (perhaps temporarily) unproductive. This contrasts with much

[1] In practice, regulators scrutinize costs even under traditional rate of regulation, providing some protection to customers.

of the existing literature which argues that the market value of a regulated firm should equal the cost of its existing assets. Our conclusion is that if regulation is to be reasonable in the presence of uncertain demand and capital prices and investment irreversibility, then the regulated firm must be allowed to derive some value from temporarily unproductive assets.

Irreversibility is intrinsic to our setup. Our analysis recognizes that assets will almost certainly be unproductive during at least part of their physical lives and that, if it is to expect future revenue to cover the cost of investment, the firm must be compensated for such periods ex ante. The regulated price that we derive does this and it implies that whenever capacity exceeds utilization the value of the regulated firm exceeds that of the green-fields replacement firm. Because temporarily unproductive assets play a similar role to investment in advance of demand (by reducing to zero the marginal cost of meeting additional demand), our analysis suggests that investment in advance for credible future demand should be allowed for in setting prices under incentive regulation.

Many authors have examined the issue of what constitutes a reasonable rate of return for a regulated firm but none, to our knowledge, have derived a closed-form solution when the firm is subject to incentive regulation with periodic resets. Leland (1974), Marshall et al. (1981), and Brennan and Schwartz (1982a, 1982b) consider the problem of calculating reasonable rates of return for firms subject to rate of return regulation. They set allowed revenues in such a way that the market value of the regulated firm equals the historical cost of the firm's assets. Kolbe and Borucki (1998) argue that, if they are to recover their cost of capital, investors need an 'insurance premium' above the cost of capital to compensate them for the risk of asset stranding induced by deregulation, but do not derive an expression for the reasonable rate of return. Hausman and Myers (2002) analyze price-setting under incentive regulation and consider the impact of irreversibility on the regulated firm. They calculate reasonable revenues using Monte Carlo simulation, but do not derive a closed-form solution for the regulated firm's reasonable rate of return. Hausman and Myers argue that the regulator has not considered the impact of investment irreversibility when calculating allowed rates of return in the US railroad industry. Hausman (1999) and Pindyck (2003) highlight similar concerns (specifically, the abandonment options which unbundling gives to the access provider's competitors) when discussing US telecommunications regulation, although they concentrate on the effect of local loop unbundling on firms' investment incentives. Salinger (1998) and Mandy and Sharkey (2003) analyze the TELRIC cost measure mandated by the FCC for the US telecommunications industry, but they ignore risk issues (and assume an exogenous cost of capital) in order to concentrate on the effects of trends in equipment costs.

Other authors have considered the more general problem of setting regulated prices when faced with non-constant demand and technology. Biglaiser and Riordan (2000) derive socially-optimal prices when future demand is uncertain and capital prices decline deterministically.

They find that the optimal output price, which must be adjusted continuously, is sensitive to demand shocks and is capped at a fixed proportion of the capital price. Beard et al. (2003) use a simple two-period model in which a regulator sets (i) the profit which a firm is allowed to earn in the event that its assets are not stranded, and (ii) the compensation which the firm will receive if its assets are stranded. The less compensation offered, the more profit must be allowed if the firm is to willingly invest in the project. The authors find that full compensation would not be offered by a welfare-maximizing regulator. Cowan (2003) shows how cost and demand shocks should affect the price which a regulated monopolist can charge. The regulator faces a trade-off between allocative efficiency and optimal risk sharing.

Our analysis extends this literature to point out the characteristics of the firm's weighted-average cost of capital and value, and the reasonable price and regulatory return, under incentive regulation when the price is fixed for periods of time.

In the next section we formally set up the model and describe the regulatory framework. Section 3 presents various measures of the cost of the regulated firm's assets, and Section 4 uses a very stylized model of incentive regulation to introduce our approach. Section 5 presents a more realistic model of price-setting, while the final section offers some concluding remarks.

2 The regulator's problem

We consider a firm which operates an infrastructure asset (referred to as 'the network' below). Let x_t equal the number of potential customers at time t, and let s_t equal the maximum number of connections (the capacity of the network), so that the number of customers actually connected to the network at time t is $\min\{x_t, s_t\}$. Shocks to x_t could arise for many reasons: population shifts can cause the number of customers wishing to connect to certain parts of the network to change, even when aggregate customer numbers do not change; the arrival of new technology may result in customers abandoning the network for others offering better services, or in new customers being attracted by new services offered by the incumbent; more generally, competition from rival providers may lead to substantial fluctuations in customer numbers. We suppose that the number of potential customers evolves according to the geometric Brownian motion

$$dx_t = \mu x_t dt + \sigma x_t d\xi_t,$$

where μ and σ are constants and ξ_t is a Wiener process. We assume that investment in new connections is irreversible.[2] The cost c_t of each new connection built at time t evolves according

[2] 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.

to the geometric Brownian motion

$$dc_t = \nu c_t dt + \phi c_t d\zeta_t,$$

where ν and ϕ are constants and ζ_t is a Wiener process satisfying $(d\xi_t)(d\zeta_t) = \rho\, dt$ for some constant $\rho \in (-1, 1)$. We price contingent claims as though x_t and c_t evolve according to the 'risk-neutral' process

$$dx_t = (\mu - \lambda)x_t dt + \sigma x_t d\xi_t, \quad dc_t = (\nu - \kappa)c_t dt + \phi c_t d\zeta_t,$$

for some constants λ and κ satisfying[3]

$$r + \lambda + \kappa > \mu + \nu + \rho\sigma\phi. \tag{1}$$

For simplicity, we assume that demand is price insensitive,[4] that connections have an infinite life, and that the network's operating costs are zero.

We suppose that the regulator forces the firm to connect any new customer who wishes to join the network. As we show in the following section, it is never optimal to invest ahead of demand. Thus, the capacity of the network is constant as long as $x_t \leq s_t$, but it is immediately raised to x_t if x_t ever climbs above s_t.[5] Figure 1 illustrates the evolution of demand and network capacity, and shows what can go wrong — the firm is forced to expand capacity as soon as there are more customers than connections; if customer numbers subsequently fall, the network is left with excess capacity.

In addition, the regulator determines the revenue which the firm is allowed to collect from its customers. The allowed revenue is to be determined by the cost structure of a hypothetical efficient replacement firm. The most important question, to which we now turn, is exactly what measure of cost is used. Sections 4 and 5 will then derive the allowed revenue.

3 Measuring the network's cost

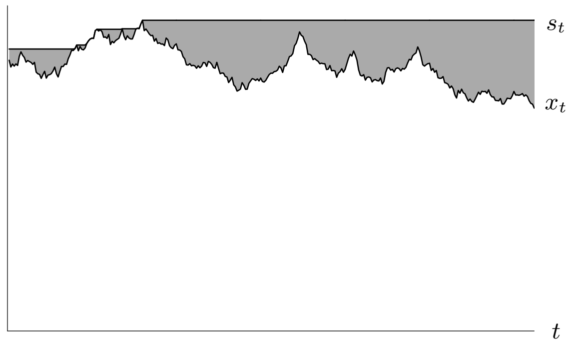
The regulated firm must invest in new connections whenever expansion of the network is required. It collects a continuous flow of revenue from customers. The regulator's task is to determine exactly how much revenue the firm can collect, a problem complicated by the fact that the firm's expansion costs are sunk. The level of revenue which is 'reasonable' depends on the

[3]We can interpret λ as the risk premium of an asset with returns which are perfectly positively correlated with changes in x_t. Thus it captures the systematic risk of shocks to customer numbers. The constant κ has a similar interpretation as the risk premium for systematic shocks to the cost of building new connections. For further discussion of risk-neutral pricing see Dixit and Pindyck (1994, Chapter 4). Condition (1) implies that the risk-adjusted discount rate for the variable cx exceeds the expected growth rate in this variable.

[4]This assumption is relaxed in Appendix B.

[5]Brennan and Schwartz (1982a) and Biglaiser and Riordan (2000) also assume that capital is a continuous variable.

Figure 1: Adjusting the capacity of the regulated network



Notes. The curve labelled x_t plots the number of potential customers as a function of time, while the curve labelled s_t plots the number of connections. As long as x_t is less than s_t the firm holds the level of s_t constant. Whenever x_t is greater than s_t the firm builds just enough new connections to ensure all customers can connect to the network. The height of the lightly-shaded region shows the number of connections which are unproductive when customer numbers fall below the network's capacity.

cost of the network, but there are many ways in which this cost can be measured. Since all investments in the network are irreversible, the cost measure of most relevance to the regulated firm is the present value of all future investment outlays that are required to meet the needs of its customers. The precise value will depend on the number of customers currently connected to the network, as well as its current capacity and the cost of new connections. We denote its value at time t by $C(x_t, s_t, c_t)$. Of more interest to a hypothetical replacement firm is the cost of replacing the network. Replicating the network at time t requires an initial outlay of $c_t s_t$. Since this results in a network with capacity of s_t, the present value of all future outlays, measured immediately after the initial investment is completed, is $C(x_t, s_t, c_t)$. Thus the present value of all costs required to replicate the network is $c_t s_t + C(x_t, s_t, c_t)$. On the other hand, replacing the network at time t with one which is optimally-configured for x_t customers requires an initial outlay of just $c_t x_t$, since it is optimal to replace the network with one having no excess capacity.[6] As the resulting network has capacity of x_t, the present value of all future outlays equals $C(x_t, x_t, c_t)$. The present value of all costs required to optimally replace the network is thus $c_t x_t + C(x_t, x_t, c_t)$.

These three measures of cost all depend on the function $C(x, s, c)$ described in the following proposition, which is proven in Appendix A.

[6]This would not necessarily be the case if capital prices were declining with the size of investment. Then it would be optimal to build excess capacity into the replacement network, although simply replicating the existing network would not necessarily be optimal. We discuss possible implications of declining capital prices in Section 6, and note that other authors have assumed constant capital prices in their models.

Proposition 1 *If the network currently has x connected customers and capacity of s, and new connections currently cost c each, then the present value of all future expenditures needed to ensure any customer who wishes to connect to the network can do so is*

$$C(x,s,c)=\begin{cases}\frac{cx}{\beta-1}\left(\frac{x}{s}\right)^{\beta-1}, & 0\le x\le s,\\ c(x-s)+\frac{cx}{\beta-1}, & s<x,\end{cases}$$

where

$$\beta=\frac{1}{2}+\frac{\lambda-\mu-\rho\sigma\phi}{\sigma^2}+\sqrt{\frac{2(r+\kappa-\nu)}{\sigma^2}+\left(\frac{1}{2}+\frac{\lambda-\mu-\rho\sigma\phi}{\sigma^2}\right)^2}. \tag{2}$$

■

Condition (1) implies that $\beta>1$, which in turn implies that

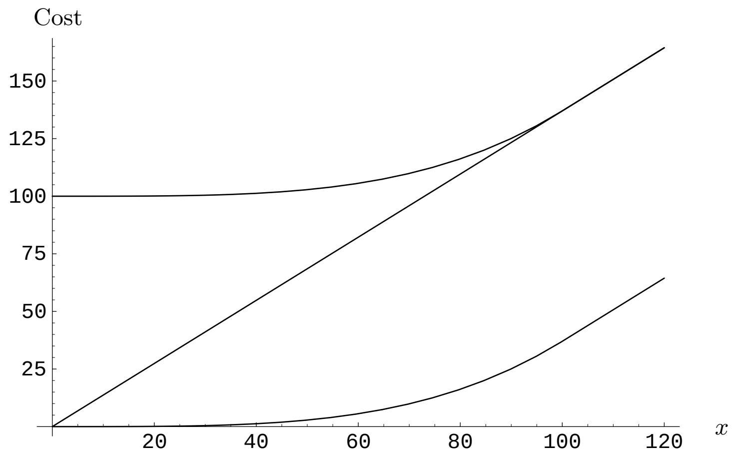
$$C(x,s,c)\le cx+C(x,x,c)\le cs+C(x,s,c).$$

That is, when all current and future costs are considered, it is cheaper to continue to operate the existing network than to replace it with one which is optimally-configured, which is cheaper than simply replicating the network.[7] Figure 2 plots the costs of the three options as a function of x. The bottom curve plots $C(x,s,c)$ as a function of x, assuming that $s=100$ connections are already in place. The straight line plots the cost of replacing the network with one which is efficiently-configured, and the top curve plots the cost of replicating the existing network.

Our measures of cost are forward-looking: they reflect both the initial outlay required to build the network as well as the stream of outflows needed to expand the network to meet future demand. Given the irreversible nature of investment in network assets, these cost measures can be significantly larger than the initial cash outlay. Proposition 1 shows that although it costs cx to build a network capable of connecting x customers, the present value of all costs equals $C(x,0,c)=\beta cx/(\beta-1)>cx$. Some typical values of the cost multiplier $\beta/(\beta-1)$ are given in Table 1. Except when customer numbers are deterministic, the multiplier exceeds unity. The extra costs do not arise just from expected growth in customer numbers, as $C(x,0,c)>cx$ even when $\mu=0$. They also arise due to volatility in customer numbers. To see why, note that any increases in customer numbers beyond x, even short-term increases, trigger investment in new connections. Because of the irreversible nature of this investment, reductions in customer numbers below x do not result in off-setting cash inflows. The asymmetry means that the present value of investment outlays exceeds the initial outlay cx. For reasonable levels of volatility, this can increase the apparent cost of building the network by one third. This means that the firm must be allowed to collect revenue with present value greater than cx if it is to be willing to build the network in the first place.

[7] Furthermore, building an additional connection costs c but (when $x<s$) only reduces the present value of all future investment expenditure by the amount $-\partial C/\partial s=c(x/s)^\beta<c$. Thus, it is never optimal to invest when the network has excess capacity.

Figure 2: Present value of all future expansion costs



Notes. The bottom curve plots $C(x, s, c)$, the present value of all future investment outlays incurred by the regulated firm, when $s = 100$. The middle curve plots $cx + C(x, x, c)$, the present value of all costs incurred by a replacement firm which replaces the network with one which is efficiently-configured. The top curve plots $cs+C(x, s, c)$, the present value of all costs incurred by a replacement firm which replicates the network. In all cases, $c = 1$, $r = 0.05$, $\lambda = \kappa = 0$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, and $\rho = 0$.

There is thus a potentially significant cost differential between the post-investment cost $C(x_t, s_t, c_t)$ that regards investment to time t as sunk, and the ex ante costs $cx_t + C(x_t, x_t, c_t)$ or $cs_t + C(x_t, s_t, c_t)$.[8] Because $C(x, s, c)$ is declining in s, a larger network requires less revenue than a smaller network to carry on in business. A firm with an existing network with capacity s, and only $x < s$ customers, will choose to stay in business if the regulator offers a revenue stream with a present value exceeding $C(x, s, c)$. However, the firm would not construct the network at the outset if it had good reason to anticipate that, once built, the network revenues would simply be sufficient to cover $C(x, s, c)$. The regulator's rules must solve this potential time-inconsistency issue and thereby induce the firm's participation under regulation.[9]

4 A simple model of incentive regulation

In their analyses of rate of return regulation, Leland (1974), Marshall et al. (1981), and Brennan and Schwartz (1982a) require that the market value of the regulated firm equals the cost of its assets. That is, they require that $R_t - C(x_t, s_t, c_t) = B_t$, where B_t denotes the historical cost of the regulated firm's assets and R_t denotes the present value of all future net revenues. We

[8] This differential would be even greater if the firm faced a sunk fixed cost of entering the industry. Such a cost would result in the top two curves in Figure 2 shifting upwards by the amount of the fixed cost.

[9] See Fudenberg and Tirole (1991, pp. 74–76) for an elaboration of time inconsistency.

Table 1: Cost multipliers

		$\sigma = 0.0$			$\sigma = 0.1$			$\sigma = 0.2$		
μ	ν	-0.02	0.00	0.02	-0.02	0.00	0.02	-0.02	0.00	0.02
-0.020		1.000	1.000	1.000	1.167	1.181	1.200	1.471	1.535	1.632
0.000		1.000	1.000	1.000	1.305	1.370	1.500	1.696	1.863	2.215
0.020		1.400	1.667	3.000	1.653	2.000	3.686	2.148	2.721	5.449

Notes. The entries in the table report the value of $\beta/(\beta-1)$, where $\beta cx/(\beta-1)$ is the present value of the investment expenditure incurred when building a network which currently has x customers. The number of customers follows a geometric Brownian motion with drift μ and volatility σ, and attracts a (systematic) risk premium of λ; the cost of connections follows a geometric Brownian motion with drift ν and volatility ϕ, and attracts a (systematic) risk premium of κ; shocks to customer numbers and the cost of a connection are uncorrelated. In all cases $\lambda = \kappa = 0$ and $r = 0.05$.

modify this to capture the key principles of incentive regulation and irreversible investment.

Definition 1 *Regulation is 'reasonable' if the present value of the firm's future net revenue stream, R_t, equals the present value of all current and future investment expenditures which a hypothetical replacement firm would incur if it were to replace the regulated firm. That is, $R_t = c_t x_t + C(x_t, x_t, c_t)$.* ■

Our definition ensures that the regulated firm's allowed revenue is determined by the cost structure of a hypothetical replacement firm which is able to build an efficiently-configured network; in particular, R_t does not depend on s_t.

We will develop a more realistic model of incentive regulation in the next section, but we begin our analysis by assuming that the firm is allowed to collect net revenue equal to the product of an allowed rate of return, $\hat{r}$, and a rate base, $c_t x_t$, that equals the cost of the assets which a hypothetical firm would build immediately if it were to replace the regulated firm. Thus, the net revenue from the firm's assets at time t is $\hat{r} c_t x_t \, dt$. Since $c_t x_t$ evolves according to the geometric Brownian motion

$$\frac{d(c_t x_t)}{c_t x_t} = (\mu + \nu - \lambda - \kappa + \rho\sigma\phi)\, dt + \sigma d\xi_t + \phi d\zeta_t,$$

it follows that the market value of this cash flow at time t equals

$$\frac{\hat{r} c_t x_t}{r + \lambda + \kappa - \mu - \nu - \rho\sigma\phi}.$$

The firm's allowed revenue is reasonable if its market value equals the value of all current and future investment expenditures which would have to be incurred by a hypothetical efficient replacement firm to serve x_t customers. This equals $\beta c_t x_t/(\beta - 1)$. Therefore[10]

[10] Appendix B extends this proposition to the case where the demand for connections is elastic. In this case the price set by the regulator affects demand for connections. While it adds a further parameter, it does not influence the qualitative conclusions of the analysis in this section.

Table 2: Determinants of allowed rates of return

		$\rho = -0.5$	$\rho = 0.0$	$\rho = 0.5$
Allowed rate of return $\hat{r}$		0.1314	0.1241	0.1171
Riskfree rate:	$\Delta r = 0.01$	0.0109	0.0109	0.0109
Expected growth rate				
Demand:	$\Delta\mu = 0.01$	−0.0072	−0.0080	−0.0086
Capital price:	$\Delta\nu = 0.01$	−0.0110	−0.0110	−0.0109
Systematic risk				
Demand:	$\Delta\lambda = 0.01$	0.0080	0.0085	0.0089
Capital price:	$\Delta\kappa = 0.01$	0.0109	0.0109	0.0109
Volatility				
Demand:	$\Delta\sigma = 0.01$	0.0035	0.0024	0.0014
Capital price:	$\Delta\phi = 0.01$	0.0008	0.0000	−0.0007

Notes. The first row reports the allowed rate of return for different values of the correlation between demand shocks and capital price shocks. The remaining rows give the change in the allowed rate of return resulting from the indicated change in the model parameter. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, and $\sigma = \phi = 0.1$.

Proposition 2 *Under reasonable incentive regulation, the allowed rate of return is*

$$\hat{r} = \left(\frac{\beta}{\beta - 1}\right)(r + \lambda + \kappa - \mu - \nu - \rho\sigma\phi) \tag{3}$$

and the value of the regulated firm equals

$$F(x_t, s_t, c_t) = \frac{\beta c_t x_t}{\beta - 1} - C(x_t, s_t, c_t).$$

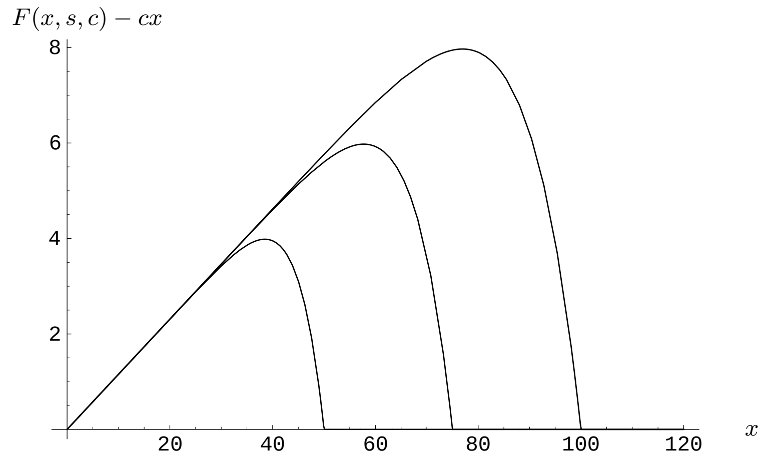
■

From equation (3), the allowed rate of return is the product of two terms. One, $r + \lambda + \kappa - \mu - \nu - \rho\sigma\phi$, is the sum of the risk free rate (r) and the risk premium for the systematic risk of the network's replacement cost ($\lambda + \kappa$), less the expected growth rate in replacement cost ($\mu + \nu + \rho\sigma\phi$). The other term is an option multiplier familiar from the real options approach to capital budgeting. This reflects the fact that when a new customer connects to the network, the firm gives this customer a valuable abandonment option: the firm makes an irreversible investment in capacity, but the customer is free to leave the network in the future.[11]

Table 2 presents information on the magnitude of the allowed rate of return under incentive regulation, and on its principal determinants. The first row of the table gives the allowed rate of return for the indicated value of ρ, showing that the premium over the risk free interest rate can be economically significant. The remaining rows of the table show the change in the allowed

[11] This asymmetry also appears when regulators impose local loop unbundling on telecommunications firms. See Hausman (1999) and Pindyck (2003) for a detailed discussion.

Figure 3: Allowed value of excess capacity



Notes. The curves plot $F(x, s, c) - cx$, the value of the firm's excess capacity (as a function of the number of customers) under incentive regulation. The top curve plots the value when $s = 100$, the middle curve when $s = 75$, and the bottom curve when $s = 50$. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, $\rho = 0$, and $c = 1$.

rate of return resulting from the indicated changes in the model's parameters. For example, if $\rho = 0$ and the expected growth rate in demand increases by one percentage point, then the allowed rate of return falls by 80 basis points. Inspection of the table reveals that the allowed rate of return is a decreasing function of the drift in both demand and capital prices, although it is more sensitive to capital price drift. Similarly, although it is an increasing function of the systematic risk of both demand and capital price shocks, the allowed rate of return is more sensitive to the systematic risk of capital price shocks. Finally, the volatility of capital price shocks has little impact on the allowed rate of return, but the volatility of demand shocks is an important determinant: even a one percentage point increase in demand volatility can lead to a 25 basis point increase in the allowed rate of return. This reflects the fact, which is clear from Table 1, that increased demand volatility increases the present value of all future investment expenditures.

Although the market value of the regulated firm equals $F(x, s, c)$, a hypothetical firm would spend just cx if it replaced the regulated firm. The difference, $F(x, s, c) - cx$, is therefore the amount by which the value of the regulated firm is allowed to exceed the replacement cost of its assets. This value is always nonnegative and is actually positive whenever the regulated firm has excess capacity (that is, $s > x$). In other words, even if the firm is only allowed to earn a 'reasonable' rate of return, its value should generally exceed that of a hypothetical efficient replacement firm. Figure 3 plots the (allowed) value of the firm's excess capacity as a function of x for three different values of s. The top curve plots the value when $s = 100$, the middle curve when $s = 75$, and the bottom curve when $s = 50$. The value of excess capacity is an increasing function of capacity (s), but a non-monotonic function of demand (x): when demand is low, there is little chance that the unproductive assets will ever be used in the future, and so the cost savings from having excess capacity are small; when demand is high, the low value of excess capacity simply reflects the small quantity of cost savings it represents.[12]

5 Setting regulated prices at discrete intervals

The previous section used a highly stylized model of incentive regulation. In this section we modify the model in two ways, aiming to make the set-up more realistic. Firstly, the regulator now sets the price for the firm's product (that is, the amount the firm can charge customers to connect to its network); secondly, it adjusts this price at discrete intervals (that is, the regulated price is held constant for finite periods of time).

Proposition 2 shows that, when the price is reset continuously, regulation is reasonable if the firm can collect revenue of $\hat{r}c_t\,dt$ from each customer. In particular, declines (respectively increases) in the capital price should be matched by declines (respectively, increases) in the firm's allowed revenue. Therefore, in this section we suppose that the regulator adjusts the output price once the capital price has changed by a significantly large margin. Specifically, we suppose that the regulator allows the firm to charge each customer connected to its network the amount $p\,dt$ at date t, where the regulated price is held fixed until the cost of connections, c_t, moves outside the band $[\underline{c}, \overline{c}]$; if the band is breached at time T, then the regulated price is immediately reset.[13] The review period is based on the capital price only, as using demand information to trigger price reviews would be more in line with rate of return regulation; we want to maintain the distinction between rate of return regulation and incentive regulation.[14]

Consistent with Definition 1, we suppose that each time the output price is set the regulator chooses a price which ensures that the present value of the future revenue stream equals $c_T x_T + C(x_T, x_T, c_T)$.[15] Thus, at each reset date the firm's allowed revenue is determined by the cost

[12]If capital prices were declining with the size of investment, then (as noted in footnote 6) it would be optimal to invest in advance of demand. The effect of such investment is similar to that of temporarily unproductive assets, that is $x_t < s_t$. This observation suggests that such investment should be reflected in the market value (and hence in the revenue) allowed the incentive-regulated firm.

[13]We also considered the possibilities that (i) revisions occur at fixed dates, and (ii) that their timing follows a Poisson process. However, in both cases there is a positive probability that capital prices will rise so high before the end of the regulatory cycle that the firm will not be able to finance investment required to connect new customers. Provided $\overline{c}$ is not set too high, this undesirable outcome can be avoided in the approach used in this section.

[14]In general, the parameters determining the behavior of demand and the capital price will enter the regulator's determination of $(\underline{\theta}, \overline{\theta})$, but in this paper we treat the regulatory policy parameters as exogenous.

[15]Brennan and Schwartz (1982a) use a similar process to determine the timing of output price changes, but use a different criterion to set the output price, namely that the market value of the firm should equal the book

structure of a hypothetical replacement firm which is able to build a network configured to reflect current demand. In particular, although the timing of the price review is triggered only by movements in the capital price, the output price chosen by the regulator also reflects the state of demand at that date. The regulatory scheme is incentive compatible in the sense that at each re-set date the firm expects to cover the cost of future investment.

The following proposition, proven in the appendix, gives the regulated price which ensures reasonable regulation.

Proposition 3 *Suppose the regulator will reset the regulated price as soon as the capital price either moves above $\overline{\theta}c$, or below $\underline{\theta}c$, where c equals the capital price at the time the regulated price was last set. Then regulation is reasonable if the regulated price is*

$$p = \frac{\beta H(\underline{\theta}, \overline{\theta})}{\beta - 1}(r + \lambda - \mu)c, \tag{4}$$

where

$$H(\underline{\theta}, \overline{\theta}) = \frac{1 - \left(\frac{(\overline{\theta})^{\delta+1} - (\underline{\theta})^{\delta+1}}{(\overline{\theta})^{\gamma+\delta} - (\underline{\theta})^{\gamma+\delta}}\right) - \left(\frac{(\overline{\theta})^{1-\gamma} - (\underline{\theta})^{1-\gamma}}{(\overline{\theta})^{-(\gamma+\delta)} - (\underline{\theta})^{-(\gamma+\delta)}}\right)}{1 - \left(\frac{(\overline{\theta})^{\delta} - (\underline{\theta})^{\delta}}{(\overline{\theta})^{\gamma+\delta} - (\underline{\theta})^{\gamma+\delta}}\right) - \left(\frac{(\overline{\theta})^{-\gamma} - (\underline{\theta})^{-\gamma}}{(\overline{\theta})^{-(\gamma+\delta)} - (\underline{\theta})^{-(\gamma+\delta)}}\right)} \tag{5}$$

and

$$\gamma = \frac{1}{2} + \frac{\kappa - \nu - \rho\sigma\phi}{\phi^2} + \sqrt{\frac{2(r + \lambda - \mu)}{\phi^2} + \left(\frac{1}{2} + \frac{\kappa - \nu - \rho\sigma\phi}{\phi^2}\right)^2}, \tag{6}$$

$$\delta = -\frac{1}{2} - \frac{\kappa - \nu - \rho\sigma\phi}{\phi^2} + \sqrt{\frac{2(r + \lambda - \mu)}{\phi^2} + \left(\frac{1}{2} + \frac{\kappa - \nu - \rho\sigma\phi}{\phi^2}\right)^2}. \tag{7}$$

The value of the firm equals

$$F(x_t, s_t, c_t) = \frac{\beta c x_t}{\beta - 1}\left(H(\underline{\theta}, \overline{\theta}) + A_1(\underline{\theta}, \overline{\theta})\left(\frac{c_t}{c}\right)^{\gamma} + A_2(\underline{\theta}, \overline{\theta})\left(\frac{c_t}{c}\right)^{-\delta}\right) - C(x_t, s_t, c_t),$$

where

$$A_1(\underline{\theta}, \overline{\theta}) = \frac{(\overline{\theta})^{\delta+1} - (\underline{\theta})^{\delta+1} - H(\underline{\theta}, \overline{\theta})\left((\overline{\theta})^{\delta} - (\underline{\theta})^{\delta}\right)}{(\overline{\theta})^{\gamma+\delta} - (\underline{\theta})^{\gamma+\delta}},$$

$$A_2(\underline{\theta}, \overline{\theta}) = \frac{(\underline{\theta})^{1-\gamma} - (\overline{\theta})^{1-\gamma} - H(\underline{\theta}, \overline{\theta})\left((\underline{\theta})^{-\gamma} - (\overline{\theta})^{-\gamma}\right)}{(\underline{\theta})^{-(\gamma+\delta)} - (\overline{\theta})^{-(\gamma+\delta)}}.$$

■

Under rate of return regulation, the firm is allowed to adjust its output price to ensure that it can recover the full cost of its investments. In particular, if demand falls, the firm is allowed to raise its output price. Customers (at least those who remain connected to the network) are thus exposed to the risk of demand shocks. In contrast, incentive regulation shields customers from this risk. However, equation (4) shows that the regulated price is proportional to the capital

value of its assets; since the resulting output price depends on the firm's actual cost structure, and not that of a hypothetical efficient firm, the approach of Brennan and Schwartz cannot be interpreted as incentive regulation.

price at the time the price is set.[16] Thus, although consumers face a constant price between revisions, they are exposed to fluctuations in capital prices in the long run. Unlike rate of return regulation, they do not bear the risk of demand shocks. In contrast, the firm is exposed to the risk that negative demand shocks will leave it unable to recover the full cost of its investments and that, in the short run, positive capital price shocks will leave it having to invest in new connections at prices which are not reflected in the regulated output price.

The multiplier $H(\underline{\theta}, \overline{\theta})$ appearing in equation (4) compensates the firm for the impact of capital price changes on the regulated price. We briefly consider four special cases. In the first case, $\underline{\theta} = \overline{\theta} = 1$; that is, the price is reset continuously. In this case, the function in equation (5) takes the value

$$H(1,1) = \frac{(\gamma - 1)(\delta + 1)}{\gamma\delta} = \frac{r + \kappa + \lambda - \mu - \nu - \rho\sigma\phi}{r + \lambda - \mu}$$

and the regulated rate of return p/c takes the value in Proposition 2. Thus, our definition of reasonable regulation is consistent with the one used in Section 4. For the remaining cases, it is useful to note that the function $H(\underline{\theta}, \overline{\theta})$ defined by (5) can be written as

$$H(\underline{\theta}, \overline{\theta}) = \frac{\overline{\theta}^{1+\delta}\left(1 - \underline{\theta}^{\delta+\gamma}\right) - \overline{\theta}^{\delta+\gamma}\left(1 - \underline{\theta}^{1+\delta}\right) - \underline{\theta}^{1+\delta}\left(1 - \underline{\theta}^{\gamma-1}\right)}{\overline{\theta}^{\delta}\left(1 - \underline{\theta}^{\delta+\gamma}\right) - \overline{\theta}^{\delta+\gamma}\left(1 - \underline{\theta}^{\delta}\right) - \underline{\theta}^{\delta}\left(1 - \underline{\theta}^{\gamma}\right)}. \tag{8}$$

In the second special case that we consider, $\underline{\theta} = 0$ and $\overline{\theta} \to \infty$; that is, the regulated price is held fixed indefinitely. From equation (8), $H(\underline{\theta}, \overline{\theta}) = 1$ and $p/c = (\beta/(\beta - 1))(r + \lambda - \mu)$. In the third special case we consider, $\underline{\theta} = 0$ and $\overline{\theta} > 1$, meaning that the regulated price is never adjusted downwards (although it will occasionally be adjusted upwards). Since $\overline{\theta} > 1$ implies that

$$H(0, \overline{\theta}) = \frac{\overline{\theta}^{\gamma} - \overline{\theta}}{\overline{\theta}^{\gamma} - 1} < 1,$$

a relatively low regulated price is reasonable when it will only be adjusted upwards in the future. In the final special case, $\underline{\theta} < 1$ and $\overline{\theta} \to \infty$, implying that the regulated price is never adjusted upwards (although it will occasionally be adjusted downwards). Since $\underline{\theta} < 1$ implies that

$$\lim_{\overline{\theta} \to \infty} H(\underline{\theta}, \overline{\theta}) = \frac{1 - \underline{\theta}^{1+\delta}}{1 - \underline{\theta}^{\delta}} > 1,$$

a relatively high regulated price is reasonable when it will only be adjusted downwards in the future.

Because of the complicated nature of the functions in Proposition 3, detailed analysis of the behavior of the regulated price requires numerical analysis. It is clear from Proposition 3 that the regulated price depends on the precise specification of the rule determining when the price is adjusted. To reduce the number of cases we need to examine, we focus attention on

[16] If there were a fixed cost of entering the industry, then the output price would have to be augmented by an additional term sufficient to ensure that the present value of the firm's revenue equals the sum of $c_T x_T + C(x_T, x_T, c_T)$ and this fixed cost. This component would be decreasing in the number of connected customers.

Table 3: Determinants of the regulated price when upward and downward revisions are equally likely

Average time between revisions (T)		0	1	5	10
Regulated price (p)		$0.1241c$	$0.1231c$	$0.1195c$	$0.1159c$
Riskfree rate:	$\Delta r = 0.01$	8.81	8.81	8.84	8.91
Expected growth rate					
Demand:	$\Delta\mu = 0.01$	−6.44	−6.44	−6.47	−6.53
Capital price:	$\Delta\nu = 0.01$	−8.83	−8.09	−5.49	−2.88
Systematic risk					
Demand:	$\Delta\lambda = 0.01$	6.87	6.88	6.91	6.98
Capital price:	$\Delta\kappa = 0.01$	8.80	8.63	8.07	7.55
Volatility					
Demand:	$\Delta\sigma = 0.01$	1.93	1.93	1.93	1.93
Capital price:	$\Delta\phi = 0.01$	0.00	−0.07	−0.32	−0.58

Notes. The first row reports the regulated price (in terms of the cost of building each connection) as a function of the average time between price revisions when upward and downward revisions are equally likely. The remaining rows give the percentage increase in the regulated price resulting from the indicated change in the model parameter, as a function of the average time between revisions. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, and $\rho = 0$.

a restricted class of rules when performing our numerical analysis. In particular, we choose values of $(\underline{\theta}, \overline{\theta})$ such that the two barriers are equally likely to be hit first; that is, upward and downward revisions of the regulated price are equally likely. In the appendix we show that this requires $(\underline{\theta}, \overline{\theta})$ to satisfy

$$2 = (\underline{\theta})^{1-2\nu/\phi^2} + (\overline{\theta})^{1-2\nu/\phi^2}. \tag{9}$$

We also show that, when $(\underline{\theta}, \overline{\theta})$ satisfy this condition the average time between price revisions is[17]

$$T = \frac{\log(\underline{\theta}\overline{\theta})}{2\nu - \phi^2}. \tag{10}$$

For a given value of T, we first solve equations (9) and (10) for $(\underline{\theta}, \overline{\theta})$, and then use Proposition 3 to calculate the regulated price.

The first row of Table 3 reports the regulated price p (in terms of the capital price c) as a function of the average time between price revisions (T) for the baseline parameters used in Section 4, revealing that more frequent revisions require a higher price. Since there is no trend in either demand or capital prices in the baseline case, this result must be due to the lower risk in the firm's net revenue when prices are held fixed for long periods of time — total net revenue varies with demand, but only responds to capital price shocks when the regulated price

[17] In the special case where $\nu = \phi^2/2$, so that the process for $\log c_t$ has zero drift, the desired behavior is achieved by setting $\underline{\theta} = e^{-\phi\sqrt{T}}$ and $\overline{\theta} = e^{\phi\sqrt{T}}$.

is revised. The remaining rows of Table 3 show how the sensitivity of the regulated price to the various model parameters varies with the frequency of price revisions. Specifically, they report the percentage increase in the regulated price for the indicated change in the model parameters. For example, increasing the riskfree interest rate by one percentage point requires raising the regulated price by 8.81 percent when the average revision interval is one year.

Higher expected demand and capital price growth each lead to lower regulated prices. Higher demand growth raises the present value of the firm's revenue stream and, by reducing the risk that investment in additional capacity will soon be stranded, lowers the present value of a hypothetical replacement firm's investment expenditure. These two effects allow a lower output price. Higher capital price growth means that the regulated price will trend upwards in the long run, thereby raising the present value of the revenue stream. Although it also raises the present value of future investment expenditure, the first effect dominates, allowing a lower output price when capital prices are expected to grow more rapidly. The role of expected demand growth is largely independent of revision frequency but, as expected, the influence of expected capital price growth is much more sensitive to revision frequency — when the regulated price is fixed for long periods of time, movements in capital prices are largely irrelevant, since they only affect the regulated price at revisions.

When the systematic risk of demand and capital price shocks is raised, the present value of future revenue and future investment expenditure both fall. Table 3 reveals that the output price rises when systematic risk rises, consistent with the present value of future revenue being more sensitive than the present value of future investment expenditure to systematic risk. The sensitivity of the regulated price to the systematic risk of demand shocks rises slightly as revision intervals become longer, while the sensitivity to the systematic risk of capital price shocks falls. This reflects the fact that with longer revision intervals the revenue stream is less sensitive to variation in the capital price.

From Table 1, the present value of all future investment expenditure increases when demand shocks are more volatile. Since demand volatility has no impact on the present value of future revenues, it follows that the regulated price must be an increasing function of demand volatility. Furthermore, since the frequency of price revisions has no impact on investment expenditure, the sensitivity of the regulated price to demand volatility is independent of revision frequency. Finally, since capital price volatility has no impact on the present value of future investment expenditure, and only a minor impact on the present value of future revenue, the regulated price is quite insensitive to changes in the capital price volatility, regardless of the frequency of revisions.

The firm's owners are also interested in the rate of return which they can expect to earn by investing in the firm as a going concern. It is easily shown that, if $F(x, s, c)$ equals the value of the firm, then this expected rate of return (the firm's weighted-average cost of capital, or

Table 4: Determinants of the firm's weighted-average cost of capital when upward and downward revisions are equally likely

		$x/s = 0.50$				$x/s = 0.75$			
Average time between revisions (T)		0	1	5	10	0	1	5	10
WACC (w)		0.1099	0.1088	0.1052	0.1016	0.1072	0.1061	0.1024	0.0988
Riskfree rate:	$\Delta r = 0.01$	0.0100	0.0099	0.0096	0.0094	0.0101	0.0100	0.0098	0.0095
Expected growth rate									
Demand:	$\Delta \mu = 0.01$	−0.0002	−0.0001	0.0001	0.0004	−0.0016	−0.0015	−0.0012	−0.0010
Capital price:	$\Delta \nu = 0.01$	−0.0000	0.0001	0.0004	0.0007	−0.0002	−0.0001	0.0002	0.0006
Systematic risk									
Demand:	$\Delta \lambda = 0.01$	0.0101	0.0100	0.0097	0.0095	0.0105	0.0104	0.0101	0.0099
Capital price:	$\Delta \kappa = 0.01$	0.0100	0.0095	0.0081	0.0070	0.0101	0.0097	0.0082	0.0071
Volatility									
Demand:	$\Delta \sigma = 0.01$	−0.0001	−0.0001	−0.0001	−0.0001	−0.0011	−0.0011	−0.0011	−0.0011
Capital price:	$\Delta \phi = 0.01$	−0.0000	−0.0000	−0.0000	−0.0001	−0.0000	−0.0000	−0.0000	−0.0001

Notes. The first row reports the firm's weighted average cost of capital (WACC) as a function of the average time between price revisions when upward and downward revisions are equally likely. The remaining rows give the change in the firm's WACC for the indicated change in the model parameter, as a function of the average time between revisions. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, and $\rho = 0$, and the WACC is calculated at the time the price is set. All results in the left half of the table are calculated assuming that $x/s = 0.50$, while $x/s = 0.75$ is used in the right half.

WACC) equals

$$w = r + \lambda \frac{x}{F}\frac{\partial F}{\partial x} + \kappa \frac{c}{F}\frac{\partial F}{\partial c}.$$

We report the firm's WACC in Table 4, using the same format as the previous tables, assuming that the regulator sets the price for network access according to equation (4). Since w is a function of x/s, we consider two representative values of this ratio.

The left panel of Table 4 reveals two main results about the firm's WACC when the network is operating with substantial excess capacity ($x/s = 0.50$). Firstly, the (systematic) risk faced by the firm's owners increases when prices are reset more frequently, as evidenced by the fact that w is a decreasing function of T. This contrasts with the finding of Brennan and Schwartz (1982a) that more frequent price setting reduces the risks faced by the owners of firms subject to rate of return regulation: such revisions allow the firm to adjust its prices in order to recover the cost of its past investments, thereby reducing the risk faced by the firm. In contrast, under incentive regulation revisions allow the regulator to adjust prices to reflect changes in capital prices, and thus pose an additional source of risk for the firm. Secondly, the only factors which affect the WACC are the riskfree interest rate and the systematic risk of demand and capital price shocks.

The influence of the first two factors is similar regardless of the frequency of price revisions. However, as we would expect, the sensitivity of the WACC to the systematic risk of capital price shocks is greatest when prices are reset more frequently (and when the firm's revenue is therefore more sensitive to capital price shocks).

The most obvious difference evident in the right panel is that the firm's WACC is lower. This reflects the fact that when there is little excess capacity, demand shocks affect both revenues and likely future investment expenditures, and the two effects offset one another; in contrast, when the network has substantial excess capacity, demand shocks affect the firm's revenue, but have little effect on the value of its future investment expenditure.[18] However, the sensitivity of the WACC to the riskfree interest rate and the premia for systematic demand and capital price risk are similar to those reported in the left panel. The drift and volatility of capital price shocks continue to have only a minor role. The main difference is that now the firm's WACC is a decreasing function of the volatility of demand shocks and, to a lesser extent, decreasing in the expected growth rate in demand.[19]

We conclude this section by considering what happens if the regulator never raises the output price; that is, if $\overline{\theta} \to \infty$. Provided $\nu < \frac{1}{2}\phi^2$, so that $\log c_t$ has negative drift, the average time between price revisions is[20]

$$T = \frac{\log \underline{\theta}}{\nu - \frac{1}{2}\phi^2}.$$

Thus, the regulator can achieve an average duration between revisions of T by setting

$$\underline{\theta} = e^{-(\frac{1}{2}\phi^2 - \nu)T}.$$

Evidence about the behavior of the regulated price in this case, following the format of Table 3, is reported in Table 5.[21] The determinants of the regulated price are similar, except that now the price is more sensitive to capital price volatility. Of course, the regulated price is higher than in Table 3, reflecting the downward trend imposed on the price by the regulator's decision not to raise the price at any date in the future. We report evidence on the firm's WACC in Table 6, using the same format as Table 4. For all but the shortest revision periods, the firm's WACC is higher than in the earlier table. In all other respects, the behavior of the WACC is similar, whether or not the output price can be revised upwards.

[18] This can be seen in Figure 2, where $C(x, s, c)$ is relatively insensitive to changes in demand when $x \ll s$.

[19] Although the sensitivities reported in the table are of similar magnitude, a one percentage point increase in the expected growth rate is a more notable event than a one percentage point increase in volatility. Thus likely variations in demand volatility would seem to have a bigger impact on the firm's WACC than likely variations in its drift.

[20] For the derivation of this result, see Dixit (1993, p. 56). The same source shows that the average time between revisions is infinite if $\nu \geq \frac{1}{2}\phi^2$.

[21] The only difference in format between the two tables is that now we reduce ν by 0.01, rather than raising it by the same amount, when considering the impact of changes in capital price drift. This is necessary to ensure that the condition $\nu < \frac{1}{2}\phi^2$ still holds.

Table 5: Determinants of the regulated price when there are no upward revisions

Average time between revisions (T)		0	1	5	10
Regulated price (p)		$0.1400c$	$0.1396c$	$0.1382c$	$0.1366c$
Riskfree rate:	$\Delta r = 0.01$	8.52	8.52	8.52	8.52
Expected growth rate					
Demand:	$\Delta\mu = 0.01$	−6.17	−6.17	−6.17	−6.18
Capital price:	$\Delta\nu = -0.01$	6.59	6.07	4.05	1.73
Systematic risk					
Demand:	$\Delta\lambda = 0.01$	6.59	6.59	6.59	6.60
Capital price:	$\Delta\kappa = 0.01$	6.59	6.59	6.59	6.59
Volatility					
Demand:	$\Delta\sigma = 0.01$	1.93	1.93	1.93	1.93
Capital price:	$\Delta\phi = 0.01$	1.93	1.87	1.66	1.41

Notes. The first row reports the regulated price (in terms of the cost of building each connection) as a function of the average time between price revisions when the regulator only revises the regulated price downwards. The remaining rows give the percentage increase in the regulated price resulting from the indicated change in the model parameter, as a function of the average time between revisions. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, and $\rho = 0$.

6 Concluding remarks

In this paper we presented a model of a firm which incorporates time, uncertain future demand and capital prices, and irreversibility in investment, and used it to analyze cost-based incentive regulation. We derived closed-form solutions for output prices which ensure that, whenever prices are reset, the present value of the firm's future net revenue stream equals the present value of the investment expenditure which a hypothetical firm would incur if it was to replace the regulated firm. Thus, the regulated price is determined by the cost structure of a hypothetical replacement firm, an approach in keeping with modern incentive regulation.

We were able to highlight some crucial differences between incentive and traditional rate of return regulation. We found that, in contrast to rate of return regulation, customers are immune to demand risk, but they are exposed to the risk of capital price fluctuations in the long run. The owners of the regulated firm bear the risk of demand shocks, as well as the risk of capital price fluctuations, and this is reflected in the firm's WACC. The frequency of price revisions affects the allocation of risk, with more frequent revisions increasing the risk faced by the firm and raising its WACC. This contrasts with the usual results of rate of return regulation, where more frequent price revisions reduce the risks faced by the firm's owners.

The market value of the regulated firm exceeds the replacement cost of its assets. The difference, which we interpret as the value of the firm's excess capacity, equals the present value of the future investment expenditure which can be avoided due to the presence of the excess

Table 6: Determinants of the firm's weighted-average cost of capital when there are no upward revisions

		$x/s = 0.50$				$x/s = 0.75$			
Average time between revisions (T)		0	1	5	10	0	1	5	10
WACC (w)		0.1099	0.1097	0.1088	0.1078	0.1072	0.1069	0.1061	0.1051
Riskfree rate:	$\Delta r = 0.01$	0.0100	0.0100	0.0099	0.0099	0.0101	0.0101	0.0101	0.0100
Expected growth rate									
Demand:	$\Delta\mu = 0.01$	−0.0002	−0.0002	−0.0002	−0.0001	−0.0016	−0.0016	−0.0015	−0.0015
Capital price:	$\Delta\nu = -0.01$	0.0000	−0.0003	−0.0017	−0.0031	0.0001	−0.0002	−0.0016	−0.0030
Systematic risk									
Demand:	$\Delta\lambda = 0.01$	0.0101	0.0101	0.0100	0.0099	0.0105	0.0105	0.0104	0.0104
Capital price:	$\Delta\kappa = 0.01$	0.0100	0.0100	0.0098	0.0096	0.0101	0.0101	0.0099	0.0097
Volatility									
Demand:	$\Delta\sigma = 0.01$	−0.0001	−0.0001	−0.0001	−0.0001	−0.0011	−0.0011	−0.0011	−0.0011
Capital price:	$\Delta\phi = 0.01$	−0.0000	−0.0000	−0.0002	−0.0003	−0.0000	−0.0000	−0.0002	−0.0003

Notes. The first row reports the firm's weighted average cost of capital (WACC) as a function of the average time between price revisions when the regulator only revises the regulated price downwards. The remaining rows give the change in the firm's WACC for the indicated change in the model parameter, as a function of the average time between revisions. Baseline parameter values are $r = 0.05$, $\lambda = \kappa = 0.03$, $\mu = \nu = 0$, $\sigma = \phi = 0.1$, and $\rho = 0$, and the WACC is calculated at the time the price is set. All results in the left half of the table are calculated assuming that $x/s = 0.50$, while $x/s = 0.75$ is used in the right half.

capacity — in effect, the firm's excess capacity provides the firm with a potentially valuable expansion option, and this option should be reflected in the market value of the firm. Note that excess capacity would only contribute to the value of the firm if there is a positive probability that it will be useful at some point in the future. While this probability might seem to be low for many unproductive assets, there are many examples of formerly unproductive assets generating revenue. Examples include electricity generators kept in mothballs for use during energy crises, and once-stranded networks rendered economically viable by new technology (such as ADSL in telecommunications networks).

In common with Brennan and Schwartz (1982a) and Biglaiser and Riordan (2000), we assume that there are no economies of scale in investment and that capital is a continuous variable. These assumptions preclude the (realistic) possibility that the regulated firm might choose to invest ahead of demand. Although the model needs to be extended to capture this choice, because temporarily unproductive assets play a similar role to assets built in advance of demand, this paper still offers insights into how such 'efficient' excess capacity should be treated by regulators. We argue that it should be treated in the same way as the excess capacity in this paper; that is, the regulated price should be set in such a way that the market value of the firm reflects the present value of future investment expenditure which can be avoided due to the presence of the excess capacity.

References

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

Beard, T. Randolph, David L. Kaserman, and John W. Mayo (2003). "Regulation, competition, and the optimal recovery of stranded costs," *International Journal of Industrial Organization* 21(6), 831–848.

Biglaiser, Gary and Michael Riordan (2000). "Dynamics of price regulation," *RAND Journal of Economics* 31(4), 744–767.

Brennan, Michael J. and Eduardo S. Schwartz (1982a). "Consistent regulatory policy under uncertainty," *Bell Journal of Economics* 13(2), 506–521.

Brennan, Michael J. and Eduardo S. Schwartz (1982b). "Regulation and corporate investment policy," *Journal of Finance* 37(2), 289–300.

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

Dixit, Avinash K. (1993). *The Art of Smooth Pasting.* Fundamentals of Pure and Applied Economics 55, Amsterdam: Harwood Academic Publishers.

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

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, MA.

Fudenberg, Drew and Jean Tirole (1991). *Game Theory*, Cambridge Mass.: MIT Press.

Hausman, Jerry (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, MA.

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.

Kolbe, A. Lawrence and Lynda S. Borucki (1998). "The impact of stranded-cost risk on required rates of return for electric utilities: Theory and an example," *Journal of Regulatory Economics* 13, 255–275.

Leland, Hayne E. (1974). "Regulation of natural monopolies and the fair rate of return," *Bell Journal of Economics and Management Science* 5(1), 3–15.

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.

Pindyck, Robert S. (2003). "Mandatory unbundling and irreversible investment in telecom networks," NBER Working Paper No. W10287.

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

Salinger, Michael A. (1998). "Regulating prices to equal forward-looking costs: Cost-based prices or price-based costs?", *Journal of Regulatory Economics* 14, 149–163.

Weisman, Dennis L. (2002). "Did the high court reach an economic low in Verizon v. FCC?" *Review of Network Economics* 1(2), 90–105.

A Proofs

Proof of Proposition 1

Let $C(x_t, s_t, c_t)$ denote the value of the regulated firm's future investment expenditures, measured at date t. If $0 \leq x < s$, then capacity remains constant over the next short time interval of length dt, during which time there are no cash flows. Thus C must satisfy

$$C(x, s, c) = e^{-r\,dt} E[C(x + dx, s, c + dc)], \quad 0 \leq x < s,$$

where r is the riskless interest rate and the expected value is calculated using the risk-neutral process for x. This implies the differential equation

$$0 = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C}{\partial x^2} + \rho\sigma\phi c x \frac{\partial^2 C}{\partial x \partial c} + \frac{1}{2}\phi^2 c^2 \frac{\partial^2 C}{\partial c^2} + (\mu - \lambda)x \frac{\partial C}{\partial x} + (\nu - \kappa)c \frac{\partial C}{\partial c} - rC, \quad 0 \leq x < s.$$

Since 0 is an absorbing boundary for x (that is, if x currently equals zero, then the network will never have any customers and there is no need for future investment), the solution must satisfy the boundary condition $C(0, s, c) = 0$ for all s and c. If $x > s$, then the network's capacity is immediately increased by $x - s$, costing $c(x - s)$ in total. Therefore

$$C(x, s, c) = C(x, x, c) + c(x - s), \quad x > s.$$

Continuity of $\partial C/\partial x$ along the boundary $x = s$ implies that

$$0 = \left.\frac{\partial C}{\partial s}\right|_{x=s} + c.$$

The structure of the problem implies that $C(x, s, c)$ is homogeneous of degree 1 in c. Therefore, we can write $C(x, s, c) = cG(x, s)$ for some function G. The differential equation for C reduces to

$$0 = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 G}{\partial x^2} + (\mu - \lambda + \rho\sigma\phi)x\frac{\partial G}{\partial x} - (r + \kappa - \nu)G, \quad 0 \le x < s,$$

and the boundary conditions reduce to $G(0, s) = 0$ and

$$0 = \left.\frac{\partial G}{\partial s}\right|_{x=s} + 1. \tag{A-1}$$

This problem has solution

$$G(x, s) = A(s)x^\beta, \tag{A-2}$$

where $\beta > 1$ is given in equation (2) and A is a function (to be determined) of the network's capacity. Substituting the function in (A-2) into condition (A-1) implies that $A'(s) = -s^{-\beta}$. Therefore

$$A(s) = \frac{s^{1-\beta}}{\beta - 1},$$

implying that the value of the firm's future investment expenditures is

$$C(x, s, c) = \begin{cases} \frac{cx}{\beta-1}\left(\frac{x}{s}\right)^{\beta-1}, & 0 \le x < s, \\ \frac{cx}{\beta-1} + c(x - s), & s < x. \end{cases}$$

Proof of Proposition 3

Let $F(x_t, c_t; p, \underline{c}, \overline{c})$ equal the present value of all future net revenues, measured at time t. If $\underline{c} < c_t < \overline{c}$ then the output price will remain at p for the next increment of time, the firm will receive cash flow of $px_t\,dt$, and F will satisfy

$$rF\,dt = px\,dt + E[dF],$$

where the expected value is calculated using the risk-neutral process. Therefore, F must satisfy the partial differential equation

$$0 = \frac{1}{2}\sigma^2 x^2 F_{xx} + \rho\sigma\phi xcF_{xc} + \frac{1}{2}\phi^2 c^2 F_{cc} + (\mu - \lambda)xF_x + (\nu - \kappa)cF_c - rF + px, \quad \underline{c} < c < \overline{c}, \tag{A-3}$$

where subscripts denote partial derivatives. Furthermore, if $c_t = \underline{c}$, then the output price is immediately reset such that the present value of all future net revenues equals $\beta c_t x_t/(\beta - 1)$. Thus

$$F(x, \underline{c}) = \frac{\beta \underline{c} x}{\beta - 1}. \tag{A-4}$$

Similar consideration of the price resetting when $c_t = \overline{c}$ shows that

$$F(x, \overline{c}) = \frac{\beta \overline{c} x}{\beta - 1}. \tag{A-5}$$

The set-up of the problem implies that F is homogeneous of degree 1 in x. Thus, we let $F(x, c) = xG(c)$ for some function G. Substituting this into equations (A-3), (A-4) and (A-5) shows that G must satisfy

$$0 = \frac{1}{2}\phi^2 c^2 G''(c) + (\nu - \kappa + \rho\sigma\phi)cG'(c) - (r + \lambda - \mu)G(c) + p, \quad \underline{c} < c < \overline{c}, \tag{A-6}$$

together with the boundary conditions

$$G(\underline{c}) = \frac{\beta \underline{c}}{\beta - 1}, \quad G(\overline{c}) = \frac{\beta \overline{c}}{\beta - 1}. \tag{A-7}$$

The system comprising equations (A-6) and (A-7) has solution

$$G(c) = \frac{p}{r + \lambda - \mu} + A_1 c^\gamma + A_2 c^{-\delta}, \tag{A-8}$$

where γ and δ are given by equations (6) and (7) respectively, and

$$\begin{aligned} A_1 &= \frac{1}{(\overline{c})^{\gamma+\delta} - (\underline{c})^{\gamma+\delta}} \left(\frac{\beta}{\beta - 1} \left((\overline{c})^{\delta+1} - (\underline{c})^{\delta+1} \right) - \frac{p}{r + \lambda - \mu} \left((\overline{c})^{\delta} - (\underline{c})^{\delta} \right) \right), \\ A_2 &= \frac{1}{(\underline{c})^{-(\gamma+\delta)} - (\overline{c})^{-(\gamma+\delta)}} \left(\frac{\beta}{\beta - 1} \left((\underline{c})^{1-\gamma} - (\overline{c})^{1-\gamma} \right) - \frac{p}{r + \lambda - \mu} \left((\underline{c})^{-\gamma} - (\overline{c})^{-\gamma} \right) \right). \end{aligned}$$

All that remains is to choose the output price p such that, at the time the price is set, the present value of the future net revenue stream equals $C(x, 0, c)$. Suppose, without loss of generality, that the price is set at time 0. Then we require p to be such that

$$F(x_0, c_0; p, \underline{c}, \overline{c}) = \frac{\beta c_0 x_0}{\beta - 1}.$$

Equivalently,

$$G(c_0) = \frac{\beta c_0}{\beta - 1}.$$

Substituting in (A-8) shows that p must satisfy

$$\frac{p}{r + \lambda - \mu} + A_1 c_0^\gamma + A_2 c_0^{-\delta} = \frac{\beta c_0}{\beta - 1}.$$

Finally, substituting in the above expressions for A_1 and A_2 and solving the resulting equation for p shows that

$$p = \frac{\beta(r + \lambda - \mu)c_0}{\beta - 1} H\left(\frac{\underline{c}}{c_0}, \frac{\overline{c}}{c_0}\right), \tag{A-9}$$

where $H(\underline{\theta}, \overline{\theta})$ is given by equation (5).

Proof of equations (9) and (10)

First, note that $\log c_t$ follows simple Brownian motion with drift $\nu - \frac{1}{2}\phi^2$ and volatility ϕ. Prices are reset when $\log c_t$ moves outside the band $[\log c_0 + \log \underline{\theta}, \log c_0 + \log \overline{\theta}]$. Using equation (6.4) of Dixit (1993), the lower barrier is reached before the upper one with probability

$$\frac{1 - \exp\left((1 - \frac{2\nu}{\phi^2}) \log \overline{\theta}\right)}{\exp\left((1 - \frac{2\nu}{\phi^2}) \log \underline{\theta}\right) - \exp\left((1 - \frac{2\nu}{\phi^2}) \log \overline{\theta}\right)}.$$

This probability equals 1/2 if and only if (9) holds.

From equation (6.8) of Dixit (1993), in general the expected time between price resettings is

$$T = \frac{\log(\overline{\theta}/\underline{\theta})}{\nu - \frac{1}{2}\phi^2} \left(\frac{\exp\left((1 - \frac{2\nu}{\phi^2}) \log \underline{\theta}\right) - 1}{\exp\left((1 - \frac{2\nu}{\phi^2}) \log \underline{\theta}\right) - \exp\left((1 - \frac{2\nu}{\phi^2}) \log \overline{\theta}\right)} + \frac{\log \underline{\theta}}{\log(\overline{\theta}/\underline{\theta})} \right).$$

If condition (9) holds, then this reduces to equation (10).

B Continuous resetting with elastic demand

Proposition 4 *Suppose that the demand for connections at date t equals $x_t = p_t^{-\varepsilon} y_t$, for some constant ε satisfying $0 < \varepsilon < 1$, where p_t is the regulated price of a connection and the risk-neutral process for the demand parameter y_t is the geometric Brownian motion*

$$dy_t = \alpha y_t dt + \psi y_t d\eta_t.$$

The cost c_t of building an additional connection at date t evolves according to the geometric Brownian motion

$$dc_t = (\nu - \kappa) c_t dt + \phi c_t d\zeta_t,$$

where $(d\eta_t)(d\zeta_t) = \rho_{cy} dt$. Then reasonable regulation results if the regulator sets the output price equal to

$$p_t = \frac{\beta c_t}{\beta - 1} \left(r - \alpha + (\varepsilon - 1) \left(\nu - \kappa + \rho_{cy}\phi\psi - \frac{1}{2}\varepsilon\phi^2 \right) \right),$$

where

$$\begin{aligned} \beta &= \frac{\frac{1}{2}\psi^2 - \alpha - \rho_{cy}\phi\psi + \varepsilon(\nu - \kappa + \frac{1}{2}\phi^2)}{\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2} \\ &\quad + \sqrt{\frac{2(r + \kappa - \nu)}{\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2} + \left(\frac{\frac{1}{2}\psi^2 - \alpha - \rho_{cy}\phi\psi + \varepsilon(\nu - \kappa + \frac{1}{2}\phi^2)}{\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2} \right)^2}. \end{aligned} \tag{B-1}$$

Proof. We show that reasonable regulation results if the regulator sets the output price equal to $p_t = Ac_t$ for some constant A. This form of output price implies that demand at date t equals $x_t = A^{-\varepsilon} c_t^{-\varepsilon} y_t$ and, in particular, evolves according to geometric Brownian motion. In fact,

$$dx_t = \left(\alpha - \varepsilon(\nu - \kappa) + \frac{1}{2}\varepsilon(\varepsilon + 1)\phi^2 - \varepsilon\rho_{cy}\phi\psi \right) x_t dt + \psi x_t d\eta_t - \varepsilon\phi x_t d\zeta_t.$$

Thus, we can use our earlier analysis, which assumed an exogenous process for demand, by noting that

$$(\mu-\lambda)x_t dt = E[dx_t] = \left(\alpha - \varepsilon(\nu-\kappa) + \frac{1}{2}\varepsilon(\varepsilon+1)\phi^2 - \varepsilon\rho_{cy}\phi\psi\right) x_t dt,$$

$$\sigma^2 x_t^2 dt = (dx_t)^2 = \left(\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2\right) x_t^2 dt,$$

and

$$\rho\phi\sigma c_t x_t dt = (dx_t)(dc_t) = \phi(\rho_{cy}\psi - \varepsilon\phi)c_t x_t dt.$$

Thus, we can use our earlier analysis by making the following substitutions:

$$\begin{aligned}
\mu-\lambda &= \alpha - \varepsilon(\nu-\kappa) + \frac{1}{2}\varepsilon(\varepsilon+1)\phi^2 - \varepsilon\rho_{cy}\phi\psi, \\
\sigma &= \left(\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2\right)^{1/2}, \\
\rho &= \frac{\rho_{cy}\psi - \varepsilon\phi}{\left(\psi^2 - 2\varepsilon\rho_{cy}\psi\phi + \varepsilon^2\phi^2\right)^{1/2}}.
\end{aligned}$$

In particular, the regulator should choose the constant A such that the present value of the firm's revenue flow equals

$$C(x_t, 0, c_t) = \frac{\beta c_t x_t}{\beta - 1},$$

where β is given by (B-1). Now, the firm's revenue flow equals $p_t x_t\, dt = A c_t x_t\, dt$, so that its present value equals

$$\frac{A c_t x_t}{r + \lambda + \kappa - \mu - \nu - \rho\sigma\phi}.$$

Thus, the regulator should set

$$A = \frac{\beta}{\beta-1}(r + \lambda + \kappa - \mu - \nu - \rho\sigma\phi),$$

or, in terms of the underlying parameters,

$$A = \frac{\beta}{\beta-1}\left(r - \alpha + (\varepsilon - 1)\left(\nu - \kappa + \rho_{cy}\phi\psi - \frac{1}{2}\varepsilon\phi^2\right)\right).$$

Setting $\varepsilon = 0$ establishes Proposition 2.