

Paperseries No. 09

Gert Brunekreeft and Jörg Borrmann

## The Effect of Monopoly Regulation on the Timing of Investment

May 2011

In Cooperation with Jacobs University Bremen



JACOBS  
UNIVERSITY

**Herausgeber/ Editor:**

Bremer **Energie** Institut 

College Ring 2 / Research V

28759 Bremen

Tel. +49 (0) 421 - 200-48 88

Fax +49 (0) 421 - 200-48 77

info@bremer-energie-institut.de

[www.bremer-energie-institut.de](http://www.bremer-energie-institut.de)

**Ansprechpartner:**

Nele Friedrichsen

Tel. +49 (0) 421 - 200-48 83

Fax +49 (0) 421 - 200-48 77

E-mail [friedrichsen@bremer-energie-institut.de](mailto:friedrichsen@bremer-energie-institut.de)

Vorgeschlagene Zitierweise/ Suggested citing:

Autor, (Jahr), "Titel", Bremen Energy Working Papers No. #, Bremer Energie Institut.

Die „Bremen Energy Working Papers“ werden vom Bremer Energie Institut in Kooperation mit der Jacobs University Bremen herausgegeben. Die Diskussionspapiere beinhalten wissenschaftliche Arbeiten mit vorläufigem Charakter, die den Diskurs befördern sollen. Die Inhalte spiegeln aber nicht notwendigerweise die Meinung des Instituts wider, sondern lediglich die des Autors. Weder das Bremer Energie Institut, noch die Jacobs University Bremen übernimmt Verantwortung für den Inhalt.

The „Bremen Energy Working Papers“ are published by the Bremer Energie Institut in cooperation with Jacobs University Bremen. The papers feature research and scholarship of a preliminary nature that should encourage dialogue and critical discussion. The content does not necessarily reflect the views of the Bremer Energie Institut or of Jacobs University Bremen and is the sole responsibility of the author. Neither the Bremer Energie Institut nor Jacobs University Bremen bear any responsibility concerning the contents.

# The Effect of Monopoly Regulation on the Timing of Investment

**Gert Brunekreeft\***  
Jacobs University Bremen  
College Ring 2  
28759 Bremen  
Germany

**Jörg Borrmann**  
University of Vienna  
Brünner Str. 72  
1210 Vienna  
Austria

May 4, 2011

## Abstract

This paper contributes a theoretical analysis of the effects of different types of regulation on the timing of monopoly investment in a setting with lumpy investment outlays. Concentrating on the case where investment increases the regulatory asset base, we distinguish between price-based regulation and cost-based regulation. Under cost-based regulation, investment triggers a change of regulated prices, whereas, under price-based regulation, investment does not affect them. To motivate investment, we focus on wear and tear leading to replacement investment and on demand growth resulting in expansion investment. Our main conclusion is that cost-based regulation accelerates investment compared to price-based regulation.

*Keywords:* Cost-based regulation, Expansion investment, Investment timing, Monopoly, Price-based regulation, Replacement investment.

*JEL classification:* D42, G00, L51

---

\*We are grateful for useful comments from Graeme Guthrie, Margarethe Rammerstorfer, David Starkie, Ingo Vogelsang and from participants at the workshop "Energy Regulation" at Tilburg University in December 2008, the workshop "Regulation, Investments and the Quality of Energy Networks" at NMA/DTE in The Hague in May 2009, the Energy Seminar, University of East Anglia, Norwich, in June 2009, the workshop "Current Topics in the Regulation of Energy and Telecommunications Markets" at Vienna University of Economics and Business in November 2009, the Business & Economics Society International Conference in Athens in July 2010, and at the workshop "Dynamic Aspects of Investment Planning and Project Evaluation", organized by FGSV, at Vienna University of Technology in October 2010.

# 1 Introduction

The early 1980s witnessed a paradigm shift in monopoly regulation, such as the regulation of energy, telecommunications, and transportation networks. Starting in the UK with the reforms of British Telecom, the regulatory model changed from traditional cost-based regulation, e.g. different forms of rate-of-return regulation, to price-based regulation, known in different variations as price caps, revenue caps, or RPI-X, and, in its extreme form, yardstick regulation (cf., e.g., Joskow, 1989; Armstrong, Cowan, Vickers, 1994; Jamasb and Pollitt, 2000). At least in some sectors, notably electricity, we appear to witness the next paradigm shift. Huge investment needs seem to trigger more cost-based, investment-enhancing components in the regulation of networks.

In this paper, we analyze the problem of cost-increasing investment. We address monopoly networks only, and we do not deal with competitive parts of these sectors. We use and define cost-based regulation, where investment triggers a change in regulated prices. In contrast, we also use and define price-based regulation where regulated prices are independent of investment. We are fully aware that these are extreme forms, and that regulatory practice is far more refined. However, the abstraction of the extremes allows us to examine the effect of the cost pass-through factor on the timing of monopoly investment.

The key advantages of price-based (or RPI-X) regulation have been well formulated by Beesley and Littlechild (1989). Price-based regulation is claimed to require less information, to allow more flexibility in the price structure enhancing welfare, and to create stronger incentives to improve productive efficiency (aka X-efficiency) than cost-based regulation. The experience with the ability of price-based regulation to improve productive efficiency is impressive (cf., e.g., Jamasb and Pollitt, 2000). However, after more than two decades of a regulation which sets strong short-run incentives to cut costs, concern rises about long-run incentives, or, in other words, incentives for adequate investment (cf. Brunekreeft and McDaniel, 2005; Vogelsang, 2009).

Much of the analytical work in the literature addresses the issue of cost-reducing investment. This type of investment is the exact aim of price-based incentive regulation. Currently, however, several monopoly networks require massive cost-increasing investment. For instance, some of the energy networks face huge investment needs which require substantial capital expenditures. The driving forces are twofold. First, energy networks tend to be subject to investment cycles. In several countries, networks tend to be old and depreciated and the investment cycles are at the foot of the hill. Second, the struggle against climate change means that network roll-out has to be adjusted. In particular, large scale integration of renewable energies requires substantial reinforcement of both transmission and distribution net-

works. Thus, the current situation faced by many regulators is large required investment to renew and reinforce networks. Therefore, in the energy sector, the main discussion is no longer on cost-reduction, but rather on incentives for cost-increasing investment.

An excellent overview of the literature on regulation and investment is provided by Guthrie (2006). This overview suggests that much of the literature concentrates on one of three problems. Some authors look at the effects of rate-of-return regulation on investment following the seminal approach of Averch and Johnson (1962). Others address the regulation of network charges in a more general setting and the effects on vertical foreclosure. This line of literature was strongly inspired by the trend towards vertical re-integration in the telecommunications sector in the United States some time around 2000. Still others analyze the short-term efficiency incentives resulting from price-based regulation.

We find only little literature on the timing of monopoly investment under regulation. An important paper is Biglaiser and Riordan (2000), who study the dynamics of price regulation for a firm adjusting to exogenous technological progress. They show that the timing of cost-reducing investment is affected by the regulatory process due to the inflexibility of depreciation rules used in practice. Another line of the literature focuses on a situation, where investment takes place under uncertainty like Dobbs (2004) and following up on that Nagel and Rammerstorfer (2009). These models rely on the real options literature. Still another line follows from a debate in Australia resulting in the concept of access holidays (cf. Gans and Williams, 1999; Gans and King, 2004). Ultimately, this line of work examines the regulatory non-commitment problem, which was explored in a game-theoretical setting by Gilbert and Newbery (1994) and also relies on uncertainty as a driver. Our paper contributes an analysis of the effects of different types of network regulation on the timing of monopoly investment under certainty. It concentrates on investment that increases the cost of the network. We formalize a simple and straightforward causal relationship: cost-based regulation accelerates large and lumpy investment as compared to price-based regulation. It is this lumpiness, i.e. investment of fixed size, which is the main driver of our main result. Therefore, if regulators are concerned about delayed network investment, they may want to introduce investment-inducing cost-based elements into the regulatory framework.

In an intertemporal model, analytically relying on Katz and Shapiro (1987), Gans and Williams (1999), Brunekreeft and Newbery (2006) and, in particular, Borrmann and Brunekreeft (2011), we analyze the behavior of a monopoly firm in different settings. We address the cases of a firm maximizing discounted social welfare, an unregulated firm maximizing discounted profits, a firm maximizing discounted profits subject to an extreme form of price-based regulation, i.e. yardstick regulation, and a firm maximizing discounted profits subject to cost-based regulation. To motivate investment,

we distinguish between two different scenarios. The first scenario is wear and tear, which is assumed to increase marginal production costs in time, while marginal costs are constant in output at any single point in time. The second scenario is demand growth, which affects the demand function. Wear and tear leads to *replacement* investment, whereas demand growth results in *expansion* investment. For the direction of the effects on the timing of investment under regulation, the difference between replacement investment and expansion investment is crucially important. We exclude the possibility of a race for investment, or, even stronger, of strategic investment to deter entry. We note that such an additional dimension is likely to affect results, but leave this for further research.

The structure of the paper is as follows. Section 2 briefly sets out the general properties of our approach and characterizes the difference between price-based regulation and cost-based regulation. Section 3 concentrates on the case of wear and tear and thereby on replacement investment. Section 4 analyzes the case of demand growth and thus concentrates on expansion investment. Section 5 concludes.

## 2 The general model

We consider a single-product monopoly firm aiming to invest in productive assets once. Investing necessitates an initial outlay,  $I$ , with  $I \in \mathbb{R}^+$ , at a single point in time without any additional outlays afterwards. The discount rate is denoted by  $r$ , with  $r \in \mathbb{R}^+$ . The firm has to make several decisions simultaneously. It has to decide which outputs to set before investment, which outputs to set after investment and when to invest in the assets. Investment is completely irreversible in the sense that there is no alternative use for the assets after investing. Investment is lumpy in the following sense. Only the investment date is a decision variable in our optimization problem (say, timing), while the choice of capacity is exogenous to the model. Capacity and technology are given. However, short-run output is optimized simultaneously.

We assume sufficiently large investment of fixed size. Therefore, investment takes place only once. This is compatible with our aim to analyze the regulation of (natural) monopoly networks where large and lumpy investment is the rule rather than the exception. The main feature of the timing problem is sufficiently large and lumpy investment. If investment in small increments is possible, the timing problem loses relevance. Instead of making a large investment every so many years, the firm will simply make a small investment every day. Obviously, timing is not much of an issue in the latter case. If investment is large, implying that investment takes place occasionally and that the investment sequence reduces until eventually only one investment remains, timing is an issue. We use the assumption that

investment takes place only once throughout this paper. This simplifies the analysis substantially. However, such an important assumption obviously needs a careful justification. In Borrmann and Brunekreeft (2011), we show formally that the analysis is similar for infinitely repeated and unrepeated investment. As expected, the length of the time intervals between different investment dates depends on the (exogenous) size of the investment. Therefore, we can be confident that the analysis in this paper also holds for infinitely repeated investment.

Thus, we consider two periods,  $i = 1, 2$ . We call the period before investment (ante-investment) period 1, and we call the period after investment (post-investment) period 2.

The objective function of the firm is either discounted social welfare, or discounted profits, either unregulated or under some specified form of regulation. The investment allows the firm to attain strictly positive discounted social welfare or strictly positive discounted profits. Thus, we distinguish between four cases using superscripts:

- $DSW$  ... discounted social welfare maximization,
- $D\Pi$  ... unregulated discounted profit maximization,
- $YR$  ... discounted profit maximization under yardstick regulation as the extreme case of price-based regulation following the seminal work of Shleifer (1985), and
- $CB$  ... discounted profit maximization under cost-based regulation, which means that the price of the good produced is allowed to increase after investment.

In our intertemporal approach, we model two investment drivers explicitly. The first investment driver is *wear and tear* leading to replacement investment. The second investment driver is *demand growth* leading to expansion investment.

In order to be able to analyze the effects of wear and tear and the effects of demand growth separately, we consider two different scenarios. In the case of wear and tear, we assume that marginal costs, which are constant in output, increase at a constant rate,  $\alpha$ , in time, with  $0 < \alpha < 1$ , and that the relationship between output and price, i.e. the demand function, does not change. Thus, in this case, there is only a driver for replacement investment, and there is no driver for expansion investment. In the case of demand growth, the relationship between output and costs, i.e. the cost function, does not change, and the quantity demanded at a given price grows at a constant rate,  $g$ , with  $0 < g < 1$ . This implies that, in the case of demand growth, there is only a driver for expansion investment, and there is no driver for replacement investment. We assume  $\alpha < r$  and  $g < r$ .

The general structure of the maximization problem, let it be either constrained in a regulated setting or unconstrained in an unregulated setting, is always as follows:

$$\max_T V(T) = \int_0^T x_1(t) e^{-rt} dt + \int_T^{T_S} x_2(t) e^{-rt} dt - Ie^{-rT}, \quad (1)$$

(possibly) subject to one or two regulatory constraints. In Eq.(1),  $V(\cdot)$ , is the objective, which is a function of  $T$ , i.e. the investment date. The functions  $x_i(\cdot)$ , which depend upon time,  $t$ , where  $i = 1, 2$  denotes the periods, will be specified for the different cases. These functions represent either social welfare or profits. In this maximization problem,  $T_S$  is either an analytical cut-off point, where a rational producer stops producing altogether, or infinity. In particular, a cut-off point is needed in the case of wear and tear, where, by assumption, at any single point in time, constant marginal costs increase in time and approach infinity, if time goes to infinity, and no further investment takes place. Therefore, we identify and substitute in each case the point where production stops.

## 2.1 The cost function and wear and tear

Production costs,  $C(\cdot, \cdot, \cdot)$ , excluding the capital costs of investment, are a function of the outputs,  $Q_1$ , before investment, of the outputs,  $Q_2$ , after investment, and of time,  $t$ . At any point in time, marginal costs are constant in  $Q_1$  and  $Q_2$ , respectively. We denote the age of the existing assets at  $t = 0$  by  $\bar{T}$ . Over time, either marginal costs increase at a constant rate,  $0 < \alpha < 1$ , due to wear and tear, as the assets of the firm get older, or the relationship between output and costs does not change, i.e.  $\alpha = 0$ . Without loss of generality, we assume fixed costs of the assets prior to investment to be negligible. The investment of  $I$  at the investment date,  $T$ , brings marginal cost back to its original level:

$$C(Q_1, Q_2, t) = \begin{cases} ce^{\alpha(t+\bar{T})}Q_1; & t < T \\ ce^{\alpha(t-T)}Q_2; & T \leq t, \end{cases} \quad (2)$$

where  $c \in \mathbb{R}^+$  and  $Q_1, Q_2, \bar{T}, T, t \in \mathbb{R}_0^+$ . This cost function exhibits (cost) economies of scale.

## 2.2 The demand function and demand growth

Inverse demand,  $P(\cdot, \cdot)$ , is a function of output,  $Q$ , and of time,  $t$ . At any point in time, demand is linear. Over time, either the quantity demanded at a given price grows at a constant rate,  $g$ , with  $0 < g < 1$ , or the relationship between output and price does not change, i.e.  $g = 0$ :

$$P(Q, t) = a - be^{-gt}Q, \quad (3)$$

where  $a, b \in \mathbb{R}^+$  and  $Q, t \in \mathbb{R}_0^+$ .



This function has the agreeable property that, at an unchanged price,  $P$ , demand grows at a steady rate  $g$ , if  $0 < g < 1$ .

### 3 Wear and tear: replacement investment

#### 3.1 Maximization of discounted social welfare and discounted profits

Building on Eq.(2) and Eq.(3) and assuming  $g = 0$ , we define social welfare,  $SW(\cdot, \cdot, \cdot)$ , for the case of wear and tear at any point in time, not taking into account investment outlays,  $I$ , as a function of the outputs,  $Q_1$ , before investment, of the outputs,  $Q_2$ , after investment, and of time,  $t$ . It is the sum of consumer surplus and profits:

$$SW(Q_1, Q_2, t) = \begin{cases} -\frac{b}{2}Q_1^2 + (a - ce^{\alpha(t+\bar{T})})Q_1; & t < T \\ -\frac{b}{2}Q_2^2 + (a - ce^{\alpha(t-T)})Q_2; & T \leq t, \end{cases} \quad (4)$$

where  $a, b, c \in \mathbb{R}^+$  and  $Q_1, Q_2, T, t \in \mathbb{R}_0^+$ . Partially differentiating the objective function with respect to  $Q_1$  and  $Q_2$  and setting the results equal to zero leads to the welfare-optimal quantities,  $Q_1^{SW}(\cdot)$  and  $Q_2^{SW}(\cdot)$ , which are functions of time,  $t$ :

$$Q_1^{SW}(t) = \frac{a - ce^{\alpha(t+\bar{T})}}{b} \quad (5)$$

and

$$Q_2^{SW}(t) = \frac{a - ce^{\alpha(t-T)}}{b}. \quad (6)$$

Denote social welfare in period 1, given the optimal quantities, by  $SW_1(\cdot)$ ; it is a function of time,  $t$ . Analogously, denote social welfare in period 2, given the optimal quantities, by  $SW_2(\cdot)$ ; it is also a function of time,  $t$ .

Alternatively, again building on Eq.(2) and Eq.(3) and assuming  $g = 0$ , we define profits,  $\Pi(\cdot, \cdot, \cdot)$ , for the case of wear and tear at any point in time, not taking into account investment outlays,  $I$ , as a function of the outputs,  $Q_1$ , before investment, of the outputs,  $Q_2$ , after investment, and of time,  $t$ :

$$\Pi(Q_1, Q_2, t) = \begin{cases} -bQ_1^2 + (a - ce^{\alpha(t+\bar{T})})Q_1; & t < T \\ -bQ_2^2 + (a - ce^{\alpha(t-T)})Q_2; & T \leq t, \end{cases} \quad (7)$$

where  $a, b, c \in \mathbb{R}^+$  and  $Q_1, Q_2, T, t \in \mathbb{R}_0^+$ . The optimal quantities,  $Q_1^{\Pi}(\cdot)$  and  $Q_2^{\Pi}(\cdot)$ , for this case, i.e. unregulated profit maximization, can easily be found. They are also functions of time,  $t$ :

$$Q_1^{\Pi}(t) = \frac{a - ce^{\alpha(t+\bar{T})}}{2b} \quad (8)$$

and

$$Q_2^\Pi(t) = \frac{a - ce^{\alpha(t-T)}}{2b}. \quad (9)$$

Denote profits in period 1, given the optimal quantities, by  $\Pi_1(\cdot)$ ; they are a function of time,  $t$ . Analogously, denote profits in period 2, given the optimal quantities, by  $\Pi_2(\cdot)$ ; they are also a function of time,  $t$ .

As noted above, our approach allows to invest only once. For the case of wear and tear, this creates a problem, if  $t$  gets large. Since marginal costs increase in time, at some point, marginal costs will be so high that a rational producer stops producing altogether. We can determine a cut-off point,  $T_S$ , beyond which no production takes place anymore, i.e.  $Q = 0$ . In particular,

$$T_S = T + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}, \quad (10)$$

where  $T$  is the investment date.<sup>1</sup> This formula applies both to a firm maximizing discounted social welfare and to a firm maximizing discounted profits.

Building on Eq.(1), we first derive the optimal investment date for the case of discounted social welfare maximization, i.e.

$$\max_T V^{DSW}(T) = \int_0^T SW_1(t) e^{-rt} dt + \int_T^{T_S} SW_2(t) e^{-rt} dt - Ie^{-rT}. \quad (11)$$

After differentiating with respect to  $T$ , setting the result equal to zero, and rearranging, we can characterize the investment date,  $T^{DSW}$ , which maximizes discounted social welfare:

$$\frac{1}{2b} \left( -c^2 e^{2\alpha(T^{DSW} + \bar{T})} + 2ace^{\alpha(T^{DSW} + \bar{T})} - \psi \right) = rI, \quad (12)$$

where  $\psi$  is given by:  $\psi = a^2 + a^2 \left( e^{-r\frac{\ln(\frac{a}{c})}{\alpha}} - 1 \right) + \frac{2acr}{\alpha-r} \left( e^{(\alpha-r)\frac{\ln(\frac{a}{c})}{\alpha}} - 1 \right) - \frac{c^2 r}{2\alpha-r} \left( e^{(2\alpha-r)\frac{\ln(\frac{a}{c})}{\alpha}} - 1 \right)$ .

Using  $e^{\alpha\frac{\ln(\frac{a}{c})}{\alpha}} = \frac{a}{c}$ ,  $e^{2\alpha\frac{\ln(\frac{a}{c})}{\alpha}} = \frac{a^2}{c^2}$ , and  $e^{-r\frac{\ln(\frac{a}{c})}{\alpha}} = \frac{1}{\left(\frac{a}{c}\right)^{\frac{r}{\alpha}}}$  we can express  $\psi$  in a more convenient way:

$$\psi = \frac{a^2 + \frac{2a^2 r}{\alpha-r} - \frac{ra^2}{2\alpha-r}}{\left(\frac{a}{c}\right)^{\frac{r}{\alpha}}} - \frac{2acr}{\alpha-r} + \frac{c^2 r}{2\alpha-r}. \quad (13)$$

---

<sup>1</sup>The necessity for determining a cut-off point is a consequence of using a two-period, one-investment model. A cut-off point is redundant in a multiple-period, sequential investment model. In Borrmann and Brunekreeft (2011), we show the similarities between the two approaches.

In Eq.(12), substitute  $T$  for  $T^{DSW}$  and denote the LHS as  $f^{DSW}(T)$ . Repeating this for unregulated discounted profit maximization gives:

$$\max_T V^{D\Pi}(T) = \int_0^T \Pi_1(t) e^{-rt} dt + \int_T^{T_S} \Pi_2(t) e^{-rt} dt - Ie^{-rT}. \quad (14)$$

Now, we can describe the investment date,  $T^{D\Pi}$ , which maximizes discounted profits:

$$\frac{1}{4b} \left( -c^2 e^{2\alpha(T^{D\Pi} + \bar{T})} + 2ace^{\alpha(T^{D\Pi} + \bar{T})} - \psi \right) = rI. \quad (15)$$

Note that, in Eq.(15),  $\psi$  is equal to  $\psi$  in Eq.(12).

In Eq.(15), substitute  $T$  for  $T^{D\Pi}$  and denote the LHS as  $f^{D\Pi}(T)$ .

Comparing discounted social welfare maximization and discounted unregulated profit maximization, it follows for the case of wear and tear that the investment date under discounted social welfare maximization is unambiguously earlier than the investment date under unregulated discounted profit maximization.

To see this, first note that the relationships  $\frac{df^{DSW}(T)}{dT} > 0$  and  $\frac{df^{D\Pi}(T)}{dT} > 0$  hold for the relevant ranges,  $a > ce^{\alpha(T + \bar{T})}$ , beyond which no consumer is willing to consume anything at the respective price. Then, note that  $f^{DSW}(T) = 2f^{D\Pi}(T)$ . Therefore,  $T^{DSW} < T^{D\Pi}$ . Note, furthermore, that, since  $\frac{df^{DSW}(T)}{dT} > 0$  and  $\frac{df^{D\Pi}(T)}{dT} > 0$  hold for the relevant ranges, and  $\frac{d(rI(T))}{dT} = 0$ , we know that the optima (in the relevant ranges) are maxima. This holds for the entire analysis in this paper. We stress that the magnitude of the effect depends critically on the size of the investment,  $I$ . As mentioned before, we concentrate on large and lumpy investment.

In this paper, we concentrate on the effect of regulation on the timing of investment. Therefore, the stated result is not our focus here. It serves as a reference point only. However, the result has been derived and discussed more formally in another paper (Borrmann and Brunekreeft, 2011). We refer you to the other paper for further details.

### 3.2 Wear and tear under regulation

We interpret the *general approach* of regulation, denoted by  $R$ , as fixing regulated prices at  $\bar{p}_1^R$  and  $\bar{p}_2^R$  in period 1 and period 2, respectively. More realistically,  $\bar{p}_1^R$  and  $\bar{p}_2^R$  can be considered upper bounds, which are binding constraints. This implies that  $\bar{p}_1^R$  and  $\bar{p}_2^R$  are always below the prices which the firm would choose left to its own devices. Furthermore, we distinguish between two different forms of regulation, i.e. price-based regulation and cost-based regulation.

We analyze *price-based regulation* by using a model of its extreme form, i.e. yardstick regulation, denoted by  $YR$ . In this special case, regulated prices of a firm are unaffected by the choice between the alternative to invest and the alternative not to invest. In other words, the regulated prices of a firm subject to yardstick regulation are independent of the underlying costs of the firm. Thus, for yardstick regulation, we assume

$$\bar{p}^{YR} = \bar{p}_1^R = \bar{p}_2^R, \quad (16)$$

where  $\bar{p}^{YR}$  is the regulated yardstick price, which depends on the costs of other firms in comparable markets.

In contrast, *cost-based regulation*, denoted by  $CB$ , means that the price of the good produced is allowed to change after investment depending on the costs incurred. Either the relationship  $\bar{p}_1^R = \bar{p}_2^R$ , or the relationship  $\bar{p}_1^R \neq \bar{p}_2^R$  holds.

The crucial difference between the two approaches is whether investment triggers a change in the regulated price or not, which is at the heart of incentive-based regulation. We stress that our model formulation is extreme as compared to regulatory practice, but this abstraction allows us to identify the cost pass-through factor on the timing of monopoly investment. Moreover, as explained in the introduction, our main interest is cost-increasing investment. This contrasts to a situation where cost-reducing efficiency improvements were the main target, while it reflects a situation with huge projected investment needs. Therefore, we focus on the case most relevant to practical purposes, where  $\bar{p}_2^R > \bar{p}_1^R$ .

This is an important, non-trivial assumption, which deserves some attention. Our aim is to analyze situations where the owners of a network intend to make a substantial investment that increases the regulatory asset base and thus requires an increase of allowed charges. In contrast, it appears that, in his literature review, Guthrie (2006) has in mind the reverse situation where the regulator reduces allowed charges, following lower costs as a result of productivity increases. The types of investment Guthrie discusses are either related to managerial effort or to process innovations. He thus concludes that it is attractive for investors to invest at the beginning of the regulatory period in order to maximize the time interval in which charges are not adjusted. Basically, this describes the incentive power of price-based regulation. However, our focus, as is clear from the set-up of the model, is on large and lumpy investment in network assets which increase cost.

Below, we develop the general approach, then we specify for price-based regulation and for cost-based regulation.

### 3.2.1 General approach

We can determine a cut-off point,  $T_S^R$ , beyond which no production takes place for the case of wear and tear, when we fix regulated prices at  $\bar{p}_1^R$  and

$\bar{p}_2^R$  in period 1 and period 2, respectively:

$$T_S^R = T + \mu^R, \quad (17)$$

where  $T$  is the investment date and  $\mu^R$  is given by  $\mu^R = \frac{\ln\left(\frac{\bar{p}_2^R}{c}\right)}{\alpha}$ . This formula applies both to a firm maximizing discounted profits subject to price-based regulation and to a firm maximizing discounted profits subject to cost-based regulation.

For given regulated prices,  $\bar{p}_i^R$ , we derive the corresponding quantities from the demand function, i.e.  $Q_i^R(\bar{p}_i^R) = \frac{a - \bar{p}_i^R}{b}$ ,  $i = 1, 2$ . The objective is to maximize discounted profits subject to the regulated prices:

$$\max_T V^R(T) = \int_0^T \Pi_1^R(t) e^{-rt} dt + \int_T^{T_S^R} \Pi_2^R(t) e^{-rt} dt - Ie^{-rT}, \quad (18)$$

where

$$\Pi_1^R(t) = \left(\bar{p}_1^R - ce^{\alpha(t+\bar{T})}\right) Q_1^R(\bar{p}_1^R), \quad (19)$$

and

$$\Pi_2^R(t) = \left(\bar{p}_2^R - ce^{\alpha(t-T^R)}\right) Q_2^R(\bar{p}_2^R). \quad (20)$$

After differentiating with respect to  $T$ , setting the result equal to zero and rearranging, we can characterize the investment date,  $T^R$ , which maximizes  $V^R(T)$ :

$$\begin{aligned} & - \left[ \bar{p}_1^R - ce^{\alpha(T^R+\bar{T})} \right] \frac{a - \bar{p}_1^R}{b} \\ & - \left[ \bar{p}_2^R \left( e^{-r\mu^R} - 1 \right) + \left( \frac{rc}{\alpha - r} \right) \left( e^{(\alpha-r)\mu^R} - 1 \right) \right] \frac{a - \bar{p}_2^R}{b} \\ & = rI. \end{aligned} \quad (21)$$

In Eq.(21), substitute  $T$  for  $T^R$  and denote the LHS as  $f^R(T)$ .

### 3.2.2 Price-based versus cost-based regulation for replacement investment

With the preliminaries above, we can now compare price-based regulation to cost-based regulation. In order to be able to do the comparison between price-based regulation and cost-based regulation, we transform price-based regulation, reflected by the yardstick price,  $\bar{p}^{YR}$ , to cost-based regulation, where the allowed prices, i.e. the ante-investment price,  $\bar{p}_1^R$ , and the post-investment price,  $\bar{p}_2^R$ , may vary between the two periods. We define

$$\bar{p}^{YR} = \gamma \bar{p}_1^R + (1 - \gamma) \bar{p}_2^R, \quad (22)$$

where  $\gamma$  is an arbitrary weighting factor;  $0 \leq \gamma \leq 1$ . Thus, our reference yardstick is a weighted average of the ante-investment price and the post-investment price. Rewriting Eq.(22) yields:

$$\bar{p}_2^R = \frac{\bar{p}^{YR} - \gamma \bar{p}_1^R}{1 - \gamma}. \quad (23)$$

The reason for doing so is to be able to compare cost-based regulation and yardstick regulation. If we use comparative statics and increase the post-investment price,  $\bar{p}_2^R$ , the above definition guarantees that the ante-investment price,  $\bar{p}_1$ , goes down, while the weighted average remains at  $\bar{p}^{YR}$ .

**Proposition 1** *For the case of wear and tear, and assuming  $\bar{p}_2^R > \bar{p}_1^R$ , cost-based regulation accelerates the optimal investment date,  $T^{CB}$ , compared to yardstick regulation where marginal revenues are non-positive. Defining  $\Delta p = \bar{p}_2^R - \bar{p}_1^R$ , we can infer that  $\frac{dT^{CB}}{d\Delta p} < 0$ . Also, the so accelerated investment date can be earlier than the investment date under discounted social welfare maximization.*

**Proof.** Reformulate the original maximization problem, as defined by Eq.(18), in more general terms:

$$\max_T V^R(T) = V_1(T) + V_2(T) - Ie^{-rT}, \quad (24)$$

where  $V_1(T) = \int_0^T \Pi_1^R(t) e^{-rt} dt$  and  $V_2(T) = \int_T^{T_S^R} \Pi_2^R(t) e^{-rt} dt$ .

Define  $\Omega(T) = \frac{dV_2(T)}{dT} e^{rT}$  and bear in mind that  $\Omega(T) < 0$ . Furthermore, note that  $\Pi_1^R(T) = \frac{dV_1(T)}{dT} e^{rT}$ . Then, we get the optimality condition:

$$-\Omega(T^R) - \Pi_1^R(T^R) = rI, \quad (25)$$

where  $e^{-rT^R}$  was deleted.

Define  $z(T^R) = -\Omega(T^R) - \Pi_1^R(T^R)$ . Now, we would like to know what happens if, starting at  $\bar{p}^{YR}$ , we increase  $\bar{p}_2^R$  and, subsequently, decrease  $\bar{p}_1^R$ . Thus, we examine  $\frac{\partial z(T^R)}{\partial \bar{p}_2^R} = -\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R} - \frac{\partial \Pi_1^R(T^R)}{\partial \bar{p}_2^R}$ . Assuming, without loss of generality,  $\gamma = 0.5$ , from Eq.(22), we know that  $\frac{d\bar{p}_1^R}{d\bar{p}_2^R} = -1$ , which leads to:

$$\frac{\partial z(T^R)}{\partial \bar{p}_2^R} = -\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R} + \frac{\partial \Pi_1^R(T^R)}{\partial \bar{p}_1^R}. \quad (26)$$

Note the minus sign in front of the first term on the RHS. From the second part in squared brackets times  $Q$  in Eq.(21), it is straightforward to deduce that

$$\Omega(T^R) = \left[ \bar{p}_2^R (e^{-r\mu^R} - 1) + \left( \frac{rc}{\alpha - r} \right) (e^{(\alpha-r)\mu^R} - 1) \right] \frac{a - \bar{p}_2^R}{b}. \quad (27)$$

Now, we need to determine  $\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R}$ :

$$\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R} = - \frac{(a - 2\bar{p}_2^R)(r - \alpha)(1 - e^{-r\mu^R}) + r(e^{r\mu^R}c - \bar{p}_2^R)e^{-r\mu^R}}{b(r - \alpha)}. \quad (28)$$

As  $e^{r\mu^R}c - \bar{p}_2^R = \left(\frac{\bar{p}_2^R}{c}\right)^{\frac{r}{\alpha}}c - \bar{p}_2^R > 0$  for  $r > \alpha$ , and since  $r > \alpha$  by assumption, we can infer that  $\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R} < 0$  for  $\bar{p}_2^R \leq \frac{a}{2}$ , i.e. where marginal revenues are non-positive.

For  $\Pi_1^R(T^R)$ , as a special case of Eq.(19), we find:

$$\frac{\partial \Pi_1^R(T^R)}{\partial \bar{p}_1^R} = \frac{a - 2\bar{p}_1^R + ce^{\alpha(T^R + \bar{T})}}{b}. \quad (29)$$

Obviously,  $\frac{\partial \Pi_1^R(T^R)}{\partial \bar{p}_1^R} > 0$  for  $\bar{p}_1^R \leq \frac{a}{2}$ , i.e. where marginal revenues are non-positive.

Substituting  $\frac{\partial \Omega(T^R)}{\partial \bar{p}_2^R}$  and  $\frac{\partial \Pi_1^R(T^R)}{\partial \bar{p}_1^R}$  into Eq.(26), we conclude that

$$\frac{\partial z(T^R)}{\partial \bar{p}_2^R} > 0, \quad (30)$$

if marginal revenues are non-positive (as a sufficiency condition).

The next step in the proof is to see that  $\frac{d\Pi_1^R(T)}{dT^R} < 0$  and  $\frac{d\Omega(T)}{dT^R} = 0$ . Therefore, given that  $\frac{\partial z(T^R)}{\partial \bar{p}_2^R} > 0$ , we find that  $T$  must go down to restore the optimality condition. This completes the proof of the first part of the proposition.

In order to prove the second part of the proposition, a numerical example suffices. Using the parameter values as above, i.e.  $a = 100$ ,  $b = 1$ ,  $c = 40$ ,  $I = 1000$ ,  $r = 0.07$ ,  $\alpha = 0.01$  as well as  $\bar{T} = 10$ , and using  $\bar{p}^{YR} = 60$ , gives  $T^{YR} \approx 7.8$ . Introducing a cost-based approach, with  $\bar{p}_1^{CB} = 58$  and  $\bar{p}_2^{CB} = 62$ , we find that  $T^{CB} \approx 2.6 < T^{DSW} \approx 7$ . ■

The sufficiency condition to derive the result, i.e. that marginal revenues are non-positive, makes perfect sense. Proposition 1 holds at least, if marginal revenues are below zero. As we are dealing with cases of binding

regulation, this is a reasonable assumption. For large values of marginal revenues, the effect reverses. Obviously, there is a level of the yardstick price beyond which an increase of  $\bar{p}_2^R$  and the subsequent decrease of  $\bar{p}_1^R$  is not useful. We dismiss these cases as irrelevant.

The intuition of this proposition is fairly straightforward. Under the type of cost-based regulation introduced above, an investment triggers higher post-investment prices, while, by mechanism, ante-investment profits are suppressed. It is thus intuitive that early investment is attractive. In other words, if quick investment has political priority, ignoring efficiency considerations, cost-based regulation is preferred over yardsticks. Note, however, that investment may also be inefficiently early.

## 4 Demand growth: expansion investment

### 4.1 General set-up, discounted social welfare and unregulated monopoly

The set-up in the case of demand growth is similar to the case of wear and tear, with two notable differences. Strictly speaking, we still need to work with a stopping point,  $T_S$ . However, in the case of demand growth, i.e. without wear and tear, marginal costs do not increase. Thus, we can simplify the analysis by substituting infinity for the endpoint. Furthermore, to have a reason to invest under demand growth at all, current capacity must be constrained. As long as capacity is not constrained, expansion investment is always unnecessary. Therefore, we assume that constrained optimized output,  $Q_1^*$ , in the ante-investment period is at the capacity constraint,  $\bar{K}$ . Expansion investment relieves the capacity constraint so that capacity is unconstrained thereafter, and optimized output will be  $Q_2^*$  in the post-investment period. Note that our problem formulation does not involve the optimal choice of capacity, but focuses exclusively on timing instead.

Using the notation as above and taking into account that  $g > 0$  and  $\alpha = 0$ , we formulate the problem of discounted social welfare maximization under demand growth as follows:

$$\max_T V^{DSW}(T) = \int_0^T SW_1(t) e^{-rt} dt + \int_T^\infty SW_2(t) e^{-rt} dt - Ie^{-rT}, \quad (31)$$

which, after optimizing for  $T$ , rearranging terms and rewriting, leads to the optimality condition for maximizing discounted social welfare,  $T^{DSW}$ :

$$\frac{(a-c)^2}{2b} e^{gT^{DSW}} - (a-c)\bar{K} + \frac{1}{2} b e^{-gT^{DSW}} \bar{K}^2 = rI. \quad (32)$$

In Eq.(32), substitute  $T$  for  $T^{DSW}$  and denote the LHS by  $h^{DSW}(T)$ .



Repeating the optimization for the case of unregulated discounted profit maximization:

$$\max_T V^{D\Pi}(T) = \int_0^T \Pi_1(t) e^{-rt} dt + \int_T^\infty \Pi_2(t) e^{-rt} dt - Ie^{-rT}. \quad (33)$$

Optimizing for  $T$ , rearranging terms and rewriting then leads to the following optimality condition, where the investment date,  $T^{D\Pi}$ , which maximizes discounted profits, is determined by:

$$\frac{(a-c)^2}{4b} e^{gT^{D\Pi}} - (a-c)\bar{K} + be^{-gT^{D\Pi}}\bar{K}^2 = rI. \quad (34)$$

In Eq.(34), substitute  $T$  for  $T^{D\Pi}$  and denote the LHS by  $h^{D\Pi}(T)$ .

Comparing these benchmark cases, it can be inferred that a private monopoly maximizing discounted profits decelerates the investment date compared to a monopoly maximizing discounted social welfare. This result is analogous to the result in the case of wear and tear. To summarize, for the case of demand growth, the optimal investment date,  $T^{DSW}$ , under discounted social welfare maximization is unambiguously earlier than the optimal investment date,  $T^{D\Pi}$ , under unregulated discounted profit-maximization, i.e.  $T^{DSW} < T^{D\Pi}$ . Since this is not our main point here, we refer you to Borrmann and Brunekreeft (2011), where this result has been derived formally and discussed in detail, for a formal proof.

## 4.2 Demand growth under regulation

### 4.2.1 General

Capacity constraints and price regulation create a tension. The market clearing prices under a capacity constraint can be higher than the allowed regulated prices, which is an impossibility in economic terms.<sup>2</sup> Our approach to address this problem is as follows. The profit of the regulated firm is determined by the regulated prices,  $\bar{p}_1^R$  and  $\bar{p}_2^R$ , while the market clearing prices,  $p_1$  and  $p_2$ , determine the quantities. These quantities are derived from the demand function at time  $t$ . By assumption,  $\bar{p}_1^R < p_1$  at  $Q_1 = \bar{K}$ , and  $\bar{p}_2^R = p_2$  at  $Q_2 = Q_2(p_2)$ . The differences between the market clearing prices and the regulated prices result in a rent which accrues to the state. This is, for instance, what happens with scarce capacity of cross-border electricity interconnectors in Europe. As a rule, scarce capacity is auctioned. The owners are not allowed to retain the auction revenue over and above

<sup>2</sup>This is a well-known problem for severely capacity-constrained airports. See, for instance, Starkie (2008).

the regulated revenue of the lines. Instead, they either lower the network charges somewhere in their network or use the excess revenue to upgrade the network and to mitigate capacity constraints. Thus, we assume:  $\Pi_1(t) = (\bar{p}_1^R - c) Q_1^R(t)$ , with  $Q_1^R(t) = \bar{K}$ . It is obvious that  $(p_1 - \bar{p}_1^R) \bar{K}$  is not part of the profit.

Maximization of the objective function,  $V^R(\cdot)$ , depending on  $T$ :

$$\max_T V^R(T) = \int_0^T (\bar{p}_1^R - c) Q_1^R(t) e^{-rt} dt + \int_T^\infty (\bar{p}_2^R - c) Q_2^R(t) e^{-rt} dt - I e^{-rT}. \quad (35)$$

Inserting the quantities,

$$Q_1^R(t) = \bar{K}, \text{ and } Q_2^R(t) = \frac{(a - \bar{p}_2^R) e^{gt}}{b}, \quad (36)$$

optimizing for  $T$ , rearranging and rewriting then gives the optimality condition describing the investment date,  $T^R$ , maximizing discounted profits under regulation with demand growth:

$$\frac{(\bar{p}_2^R - c) (a - \bar{p}_2^R) e^{gT^R}}{b} - (\bar{p}_1^R - c) \bar{K} = rI. \quad (37)$$

In Eq.(37), substitute  $T$  for  $T^R$  and denote the LHS by  $h^R(T)$ .

#### 4.2.2 Price-based versus cost-based regulation under expansion investment

Using the mechanism to compare price-based regulation with cost-based regulation as defined in Section 3.2.2, with  $\bar{p}^{YR} = \gamma \bar{p}_1^R + (1 - \gamma) \bar{p}_2^R$  and  $\Delta p = \bar{p}_2^R - \bar{p}_1^R$ , we are able to state the following proposition.

**Proposition 2** *For the case of demand growth, assuming  $\bar{p}_2^R > \bar{p}_1^R$ , cost-based regulation accelerates the investment date for  $\bar{K} > 0$  compared to price-based regulation, whereas the investment dates for cost-based regulation and price-based regulation are equal for  $\bar{K} = 0$ . The following relationship holds:  $\frac{dT^{CB}}{d\Delta p} < 0$ .*

**Proof.** This result is similar to the first part of Proposition 1. Building on the formulation of the objective function in Eq.(35), we find immediately that

$$\Pi_2^R(T^{CB}) - \Pi_1^R(T^{CB}) = rI. \quad (38)$$

We define  $Z(T^{CB}) = \Pi_2^R(T^{CB}) - \Pi_1^R(T^{CB})$ . Now, it is obvious that  $\frac{\partial Z(T^{CB})}{\partial \bar{p}_2^R} = \frac{\partial \Pi_2^R(T^{CB})}{\partial \bar{p}_2^R} - \frac{\partial \Pi_1^R(T^{CB})}{\partial \bar{p}_1^R} \frac{d\bar{p}_1^R}{d\bar{p}_2^R}$ . As in the Proof of Proposition 1, we

assume, without loss of generality, that  $\frac{d\bar{p}_1^R}{d\bar{p}_2^R} = -1$ . Since  $\frac{\partial \Pi_2^R(T^{CB})}{\partial \bar{p}_2^R} > 0$  and  $\frac{\partial \Pi_1^R(T^{CB})}{\partial \bar{p}_1^R} > 0$ , it can be easily seen that  $\frac{\partial \Pi_1^R(T^{CB})}{\partial \bar{p}_2^R} < 0$ . Thus, for a given value of  $rI$ , we get  $\frac{\partial Z(T^{CB})}{\partial \bar{p}_2^R} > 0$ , which, in order to restore the optimality condition, implies that the optimal investment date needs to go down, which in turn implies  $T^{CB} < T^{YR}$ . ■

Also, we find that  $T^{CB} \leq T^{DSW}$ . In words, cost-based regulation can both accelerate and decelerate expansion investment compared to the socially optimal outcome. Indeed, it is quite likely that the timing of expansion investment under cost-based regulation is decelerated compared to discounted social welfare maximization. Nevertheless, cost-based regulation can also accelerate the investment date compared to the socially optimal investment date. This requires a sufficiently low ante-investment price, a sufficiently high post-investment price, and a sufficiently high capacity constraint.

We show this by a numerical example. Take the following parameter values:  $a = 100$ ,  $b = 1$ ,  $c = 20$ ,  $I = 25,000$ ,  $r = 0.07$ ,  $g = 0.05$ ,  $\bar{T} = 10$ , and  $\bar{K} = 30$ . Use  $\bar{p}_1^R = 20$ , and  $\bar{p}_2^R = 60$ . This gives  $T^{DSW} \approx 3.27$  and  $T^{CB} \approx 1.79$ , and therefore  $T^{CB} < T^{DSW}$ . A driver for the acceleration effect of regulation on expansion investment is that investment reduces ante-investment profits in case of an already existing strictly positive capacity constraint, i.e.  $\bar{K} > 0$ .

**Proposition 3** *For the case of demand growth and for  $\bar{K} = 0$ , we find  $\frac{dT^{CB}}{d\bar{p}_2^R} < 0$  and  $T^{CB}(= T^{YR}) \geq T^{DII} > T^{DSW}$ .*

In words, if we assume that there is no capacity before investment (green field), then a higher allowed price unambiguously accelerates the investment date. The investment date of a regulated monopoly maximizing discounted profits is always later than (or equal to) the investment date of an unregulated monopoly maximizing discounted profits, which is always later than the socially optimal date.

Above, we touched upon the different cases of a capacity constraint and the practical relevance of these cases. Basically, we need to distinguish between two extremes. First, there is the case of an emerging capacity constraint and a subsequent genuine capacity expansion. For this case, we assume, admittedly somewhat extreme, that charges always apply to the entire capacity, i.e. they apply to existing and new assets in the same way. This implies that new investment has an effect on the profitability of the existing assets. Second, if, alternatively, the existing assets can be priced without changing regulated charges, while new assets are priced differently, then the link between existing and expansion assets is broken and the assumption is analytically equivalent to the case where  $\bar{K} = 0$ .

The case of  $\bar{K} = 0$  is actually relevant and realistic, and it has a strong appeal for two reasons. First, in many cases of large new investment that can, in regulatory terms, be isolated from other parts of a firm, the analytical setting would be just that. Clear cases are new product innovations, where there are no old or existing assets. Moreover, for instance, big electricity interconnectors or big gas pipelines, will typically qualify as stand-alone investments and can well be regulated in isolation. Therefore, these cases are analytically equivalent to  $\bar{K} = 0$ . Second, the assumption that old and new assets are always characterized by the same charges may not always apply. In particular, in many cases, the use of infrastructure may be arranged with initial upfront connection charges, or, in even more cases, infrastructure use might be arranged by long-term contracts which may be insulated against capacity shortages and expansions. Moreover, regulators, faced with the threat of low investment and capacity shortages, now tend to work with rate-of-return top-ups for desired new investment (or, as it was phrased in the United States, with rate-of-return "adders"). In effect, regulators will grant higher rates of return for additional investment, which breaks the link between existing and new assets and is therefore analytically equivalent to the case where  $\bar{K} = 0$ . The propositions above suggest that this policy will indeed accelerate expansion investment.

For the case of  $\bar{K} = 0$ , the effects on timing are unambiguous, as claimed in Proposition 3. For expansion investment with  $\bar{K} = 0$ , the investment date of a regulated monopoly maximizing discounted profits is always later than (or equal to) the investment date of an unregulated monopoly maximizing discounted profits, which, in turn, is always later than the socially optimal investment date.

## 5 Concluding remarks

This paper deals with a topical problem in the area of monopoly regulation. By a monopoly, we mean a natural monopoly with sunk costs due to infrastructure investment, e.g. a transmission network or a distribution network for electricity or gas.

Many regulators around the globe are concerned about what seem to be low investment activities in these physical networks. First, many network assets are aging and need to be replaced. Second, there is skepticism regarding private incentives to maintain the quality of the network. This skepticism is especially relevant to the relationship between price-based regulation and investment in quality, starting with the discussion on the seminal work of Spence (1975). Third, frequently, regulators are actively promoting expansion investment of the network.

Our paper contributes to the theoretical literature by exploring the relationship between different types of regulation and the timing of monopoly

investment. We examine the differences between price-based regulation and cost-based regulation. We use and define cost-based regulation, where investment triggers a change in regulated prices. In contrast, we use and define price-based regulation where regulated prices are independent of investment. We are fully aware that these are extreme forms, and that regulatory practice is far more refined. However, the abstraction of the extremes allows us to identify the effect of the cost pass-through factor on the timing of monopoly investment. The paper studies large and lumpy investment outlays of a fixed nature, so that investment timing is an issue at all. Moreover, reflecting practical relevance, the paper focuses on cost-increasing investment. This is an important assumption, as the focus of the literature on investment and price-based (or incentive-based) regulation is on cost-reducing investment. We distinguish two intertemporal effects which justify investments: wear and tear, which causes replacement investment, and demand growth leading to expansion investment.

Our main message is that cost-based regulation, including rate-of-return adders and top-ups, leads to an earlier investment date than price-based regulation. However, as far as the efficiency of investment timing is important, details matter. Especially for replacement investment, a cost-based approach can quite easily accelerate the privately optimal investment date inefficiently fast. In other words, the privately optimal investment date can be too early. For expansion investment, it is very unlikely that the investment date is ever inefficiently early. In general, we conclude that, when timely investment is the regulator's prime objective and efficiency considerations are only of secondary importance, cost-based regulation for new investment is preferable over price-based regulation.

Our approach assumes a monopoly situation, and it neither presumes a race for investment nor tendering. As an issue for further research, we note that timing considerations change, if we allow a race for investment. In particular, a race for investment, if feasible at all, will accelerate the optimal investment date, as compared to the monopoly case. Yet, in many real-world situations in network industries, a race for investment is difficult to imagine. Where it is feasible, it is likely to be ineffective, or it may generate other problems. In particular, we think of merchant investors in high-voltage transmission networks (cf. Brunekreeft, 2004, 2005). Alternatively, tendering of the investment opportunity might be an option. Although against a slightly different background, this option was discussed as an option in EU legislation for transmission networks and, more generally speaking, there appears to be a development towards more decentralized investment models.

## References

- [1] Armstrong, M., Cowan, S. and Vickers, J. (1994), *Regulatory Reform: Economic Analysis and British Experience*, MIT Press: Cambridge, Mass.
- [2] Averch, H. and Johnson, L.L. (1962), Behavior of the firm under regulatory constraint, *American Economic Review*, **52** (5), 1052-1069.
- [3] Beesley, M.E. and Littlechild, S.C. (1989), The regulation of privatized monopolies in the United Kingdom, *RAND Journal of Economics*, **20** (3), 454-472.
- [4] Biglaiser, G. and Riordan, M. (2000), Dynamics of price regulation, *RAND Journal of Economics*, **31** (4), 744-767.
- [5] Borrmann, J. and Brunekreeft, G. (2011), The timing of repeated and unrepeated monopoly investment under wear and tear and demand growth, *Bremen Energy Working Papers*, No. 8, May, Jacobs University Bremen.
- [6] Brunekreeft, G. (2004), Market-based investment in electricity transmission networks: Controllable flow, *Utilities Policy*, **12** (4), 269-281.
- [7] Brunekreeft, G. (2005), Regulatory issues in merchant transmission investment, *Utilities Policy*, **13** (2), 175-186.
- [8] Brunekreeft, G. and McDaniel, T.M. (2005), Policy uncertainty and supply adequacy in electric power markets, *Oxford Review of Economic Policy*, **21** (1), 111-127.
- [9] Brunekreeft, G. and Newbery, D.N. (2006), Should merchant transmission investment be subject to a must-offer provision?, *Journal of Regulatory Economics*, **30** (3), 233-260.
- [10] Dobbs, I.M. (2004), Intertemporal price cap regulation under uncertainty, *Economic Journal*, **114** (495), 421-440.
- [11] Gans, J.S. and King, S.P. (2004), Access holidays and the timing of infrastructure investment, *Economic Record*, **80** (248), 89-100.
- [12] Gans, J.S. and Williams, P.L. (1999), Access regulation and the timing of infrastructure investment, *Economic Record*, **75** (229), 127-37.
- [13] Gilbert, R. and Newbery, D.N. (1994), The dynamic efficiency of regulatory constitutions, *RAND Journal of Economics*, **25** (4), 538-554.
- [14] Guthrie, G. (2006), Regulating infrastructure: The Impact on risk and investment, *Journal of Economic Literature*, **44** (4), 925-972.

- [15] Jamasb, T. and Pollitt, M. (2000), Benchmarking and regulation of electricity transmission and distribution utilities: Lessons from international experience, Cambridge Working Paper Series in Economics, 0101, University of Cambridge.
- [16] Joskow, P.L. (1989), Regulatory failure, regulatory reform, and structural change in the electric power industry, Working Papers from Massachusetts Institute of Technology (MIT), Department of Economics, 516.
- [17] Katz, M.L. and Shapiro, C. (1987), R&D rivalry with licensing or imitation, *American Economic Review*, **77** (3), 402-420.
- [18] Nagel, T. and Rammerstorfer, M. (2009), Modeling investment behavior under price cap regulation, *Central European Journal of Operations Research*, **17** (2), 111-129.
- [19] Shleifer, A. (1985), A theory of yardstick competition, *RAND Journal of Economics*, **16** (3), 319-327.
- [20] Spence, A.M. (1975), Monopoly, quality, and regulation, *Bell Journal of Economics*, **6** (2), 417-429.
- [21] Starkie, D. (2008), *Aviation Markets: Studies in Competition and Regulatory Reform*, Ashgate: Aldershot, Hampshire.
- [22] Vogelsang, I. (2009), Incentive regulation, investments and technological change, mimeo, Boston University.