Bremer Energie Institut

Bremen Energy Working Papers

Paperseries No. 01

Gert Brunekreeft and Jörg Borrmann

The Effect of Monopoly Regulation on the Timing of Investment.

Febuary 2010

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

Ansprechpartner:

Katy Jahnke

Tel. +49 (0) 421 - 200-48 91

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

E-mail jahnke@bremer-energie-institut.de

Vorgeschlagene Zitierweise/ Suggested citing:
Autor, (Jahr), "Titel", Bremen Energy Working Papers No. #, Bremer Energie Institut.

The Effect of Monopoly Regulation on the Timing of Investment

Gert Brunekreeft*
Jacobs University Bremen
College Ring 2
28759 Bremen
Germany
g.brunekreeft@jacobs-university.de

Jörg Borrmann
Department of Business Administration
University of Vienna
Brünner Str. 72
1210 Vienna
Austria
joerg.borrmann@univie.ac.at

January, 2010
Bremen Energy Working Papers 01

Abstract

This paper contributes a theoretical analysis of the effects of regulation on the timing of monopoly investment under certainty in a setting with lumpy investment outlays. We distinguish between price-based regulation and cost-based regulation. To motivate investment, we focus on wear and tear leading to replacement investment and on demand growth resulting in expansion investment. For replacement investment, price-based regulation may work just fine, if properly applied, but it does not work well for expansion investment. Cost-based regulation accelerates investment compared to price-based regulation, but this may not always be efficient.

*The authors 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, and the workshop "Current Topics in the Regulation of Energy and Telecommunications Markets" at Vienna University of Economics and Business in November 2009.

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., Armstrong, Cowan, Vickers, 1994; Jamasb and Pollitt, 2000).

The main drivers for change have been well formulated by Beesley and Littlechild (1989). Price-based regulation is claimed to require less information, as it relies less strongly on cost 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 latter point has received most attention. In the era of liberalization, the companies to be regulated were regarded as X-inefficient, and the scope for improvement was expected to be considerable. The incentive structure is straightforward, since price-based regulation delinks regulated prices from underlying actual costs. Therefore, if the regulated firm manages to reduce costs beyond the expectations of the regulator, it does not have to reduce prices and can retain the efficiency improvement as profits.

The similarities and differences between cost-based regulation and price-based regulation have been discussed extensively (see, in particular, Joskow, 1989). To put it in a nutshell, they can be attributed to the *regulatory lag*. The regulatory lag is the period in which the regulator does not change the rules that determine allowed prices or revenues.[1] If the regulatory lag is short and endogeneous, we tend to think of cost-based regulation, whereas, if the regulatory lag is long and exogenous, we tend to call a specific regulation price-based. Thinking in terms of polar cases, under cost-based regulation, allowed revenues vary with the firm's underlying costs in order to retain a reasonable fixed profit margin, whereas, under price-based regulation, allowed revenues or prices are independent of the firm's own costs. They are determined by something external instead, for example a weighted average of the costs of other firms in the same industry, which leads to the regulatory practice of benchmarking.

The experience with the ability of incentive regulation to improve productive efficiency is impressive (cf., e.g., Jamasb and Pollitt, 2000). However, after two decades of a regulation which sets strong short-run incentives to cut costs, concern starts to arise about long-run incentives, or, in other words, incentives for adequate investment (cf. Brunekreeft and McDaniel, 2005; Vogelsang, 2009). Thus, it is not by coincidence that the discussion about the latest regulatory periods in energy networks in the UK is dominated by investment needs. The regulator, Ofgem, takes pride in having allowed investment budgets of more than £12 billion in total in the course of 5 years. More importantly, Ofgem introduced a menu of sliding scales for the regulation of the investment needs of

[1] Note that, within the control period, prices and revenues as such can change, but following predetermined rules set by the regulator.

electricity distribution networks to address the information asymmetry between firms and the regulator (cf. Joskow, 2006; Brunekreeft, 2009). Furthermore, we observe similar concerns about efficient infrastructure investment incentives under regulation in other countries as well. For instance, regulators in Australia, in the Netherlands, in New Zealand, Norway, and not the least in the United States, work actively on this issue.

The theoretical literature on the relationship between regulation and efficient investment incentives is surprisingly thin and is only just emerging. A recent overview of the literature on regulation and investment is provided by Guthrie (2006). This overview leaves the impression that much of the literature concentrates on one of three problems. Many 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 inspired strongly 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 incentives of and the price structures resulting from price-based regulation models focussing on the reduction of costs, whereas the effects of price-based regulation models on investment incentives, i.e. the long-term incentives, were short of attention for a long time; this literature has started to develop only recently.

In particular, we find very little on the *timing* of monopoly investment under regulation. One line going in this direction is Dobbs (2004) and following up on that Nagel and Rammerstorfer (2009), where investment takes place under uncertainty. These models rely on the real options literature. Another line follows from a debate in Australia resulting in the concept of access holidays (cf. Gans and Willams 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 regulation on the timing of monopoly investment under certainty. 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 (2009),[2] 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: 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, and demand growth, which affects the demand function. Wear and tear leads to *replacement* investment,

[2] At the moment of writing the manuscript, Borrmann and Brunekreeft (2009) is under revision. The mimeo can be received from the authors upon request.

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 first 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, when to invest in the assets and how long to use them, i.e. how long to produce. 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 quantity choice (say, capacity) and the quality choice (say, technology) are exogenous to the model. Capacity and technology are given. Moreover, we do not impose a priori restrictions on the size of the inital outlay, I, although, implicitly, the key driver of the timing problem is sufficiently large investment. If investment in small increments is possible, the timing problem becomes trivial and loses relevance. However, if investment is large, implying that investment takes place occasionally and the investment sequence reduces until eventually only one investment remains, timing is an issue. Then, different optimization approaches under varying constraints lead to different results. Therefore, we assume sufficiently large investment outlays, I, so that investment takes place only once. This is compatible with our aim to analyze the regulation of (natural) monopoly where large lumpy investment is the rule rather than exception. Thus, we consider two periods, $i = 1, 2$, only. 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:

SW.....discounted social welfare maximization,

Π........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 dynamic 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, α, 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:

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

(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 an analytical cut-off point, where a rational producer stops producing altogether. This is particularly important for 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. We refer to Borrmann and Brunekreeft (2009) for a more detailed analysis.

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. The age of the existing assets at $t = 0$ is denoted by $\overline{T}$. Over time, either the relationship between output and costs does not change, i.e. $\alpha = 0$, or marginal costs increase at a constant rate, $0 < \alpha < 1$, due to wear and tear, as the assets of the firm get older. Without loss of generality, we neglect any fixed costs of the assets prior to investment. The investment of I at the investment date, T, brings marginal cost back to its original level:

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

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

2.2 The demand function and demand growth

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

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

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

3 Wear and tear: replacement investment

3.1 Maximization of social welfare and profits

In this section, we use the formulation and derivation in Borrmann and Brunekreeft (2009). Building on Eq.(2) and Eq.(3) and assuming $g = 0$, we define social welfare, $SW\left(\cdot,\cdot,\cdot\right)$, 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:

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

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, $Q_1^{SW}\left(\cdot\right)$ and $Q_2^{SW}\left(\cdot\right)$, which are functions of time, t:

$$Q_1^{SW}\left(t\right) = \frac{a - ce^{\alpha\left(t+\overline{T}\right)}}{b} \tag{5}$$

and

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

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:

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

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:

$$Q_1^{\Pi}(t) = \frac{a - ce^{\alpha\left(t+\overline{T}\right)}}{2b} \tag{8}$$

and

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

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. As explained in detail in Borrmann and Brunekreeft (2009), we can determine a cut-off point, T_S, beyond which no production takes place anymore, i.e. $Q = 0$. In particular,

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

where T is the investment date. This formula applies both to a firm maximizing discounted social welfare and to a firm maximizing discounted profits.

We first derive the optimal investment date for the case of discounted social welfare maximization, i.e.

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

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

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

where ψ is given by: $\psi =$

$$a^2+a^2\left(e^{-r\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right)+\frac{2acr}{\alpha-r}\left(e^{(\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right)-\frac{c^2r}{2\alpha-r}\left(e^{(2\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right).$$

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

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

In Eq.(12), substitute T for T^{SW} and denote the LHS as $f^{SW}(T)$.

Repeating this for unregulated discounted profit maximization gives:

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

Now, we can describe the investment date, T^{Π}, which maximizes discounted profits:

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

Note that, in Eq.(15), ψ is equivalent to ψ in Eq.(12).

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

Comparing discounted social welfare maximization and discounted unregulated profit maximization then gives the following proposition.

Proposition 1 *For the case of wear and tear, the investment date under discounted social welfare maximization is unambiguously earlier than the investment date under unregulated discounted profit-maximization:* $T^{SW}<T^{\Pi}$.

Proof. First, the relationships $\frac{df^{SW}(T)}{dT}>0$ and $\frac{df^{\Pi}(T)}{dT}>0$ hold for the relevant ranges, $a>ce^{\alpha\left(T+\overline{T}\right)}$, beyond which no consumer is willing to consume anything at the respective price. Second, $f^{SW}(T)=2f^{\Pi}(T)$. ■

This result has already been derived and explained in Borrmann and Brunekreeft (2009). Describing a dynamic inefficiency in a general way, it can be regarded as fundamental. An unregulated monopoly that maximizes discounted profits inefficiently decelerates the investment date compared to a monopoly that maximizes discounted social welfare. In this paper, we do the proofs not by solving for the investment date, but rather by comparing the f-functions. Using the technique to proof Proposition 1, our two steps suffice.

Since we use this type of analysis throughout the paper, we have to explain the idea more carefully. Consider Figure 1 below, which plots f^{SW} and f^{Π} as functions of T. Also, it plots the RHS of the optimality conditions: rI. The optimal investment dates are to be found at T^{SW} and T^{Π}, i.e. at the points of intersection, where $f^{SW}(T) = rI$ and $f^{\Pi}(T) = rI$, respectively. In this particular case, if we can show that $f^{SW}(T)$ and $f^{\Pi}(T)$ are strictly monotonically increasing in T in the relevant range and that $f^{SW}(T) > f^{\Pi}(T)$, then the point of intersection, where $f^{SW}(T) = rI$, must be to the left of the point of intersection, where $f^{\Pi}(T) = rI$, and therefore $T^{SW} < T^{\Pi}$. In the proofs and analyses below, we will argue along these lines throughout.

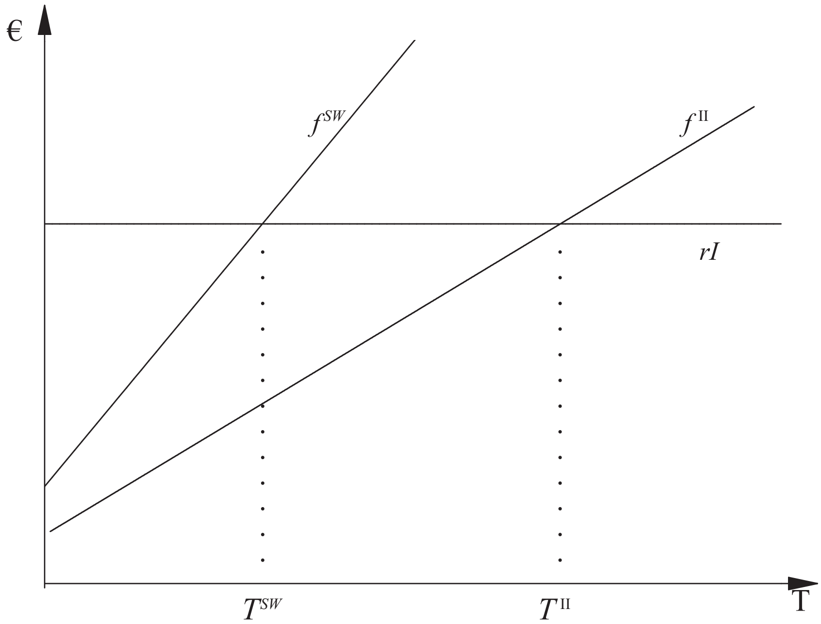

Figure 1: Illustration of f^{SW}, f^{Π} and rI *as functions of* T

3.2 Wear and tear under regulation

We interpret the *general approach* of regulation, denoted by R, as fixing regulated prices at $\overline{p}_1^R$ and $\overline{p}_2^R$ in period 1 and period 2, respectively. More realistically, $\overline{p}_1^R$ and $\overline{p}_2^R$ can be considered upper bounds, which are binding constraints. This implies that $\overline{p}_1^R$ and $\overline{p}_2^R$ 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 most extreme form, i.e. yardstick regulation, denoted by YR. In this special case, regulated prices 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

$$\overline{p}^{YR} = \overline{p}_1^R = \overline{p}_2^R, \tag{16}$$

where $\overline{p}^{YR}$ 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 $\overline{p}_1^R = \overline{p}_2^R$, or the relationship $\overline{p}_1^R \neq \overline{p}_2^R$ holds. In particular, we focus on the case most relevant for practical purposes, where $\overline{p}_2^R > \overline{p}_1^R$.

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

3.2.1 General approach

Following Borrmann and Brunekreeft (2009), we can determine a cut-off point, T_S^R, beyond which no production takes place, for the case of wear, when we fix regulated prices at $\overline{p}_1^R$ and $\overline{p}_2^R$ in period 1 and period 2, respectively:

$$T_S^R = T + \mu^R, \tag{17}$$

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

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

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

where

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

and

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

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

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

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

3.2.2 Replacement investment and price-based regulation

By substituting $\overline{p}^{YR} = \overline{p}_1^R = \overline{p}_2^R$ into $f^R(T)$ and calling this $f^{YR}(T)$, we can now characterize the outcome for the most extreme form of price-based regulation, i.e. yardstick regulation. With this, we state Proposition 2.

Proposition 2 *For the case of wear and tear and yardstick regulation, a lower yardstick price accelerates the optimal investment date, T^{YR}, i.e. $\frac{\partial T^{YR}}{\partial \overline{p}^{YR}} > 0$ in the relevant range, where the optimality condition is fulfilled. This result holds provided that the regulated price still allows full cost recovery of the investment.*

Proof. Using the optimality condition, i.e. $f^{YR}(T) = rI$, we get:

$$\left(-\frac{a-\overline{p}^{YR}}{b}\right)\left[\overline{p}^{YR}e^{-r\mu^{YR}} - ce^{\alpha\left(T^{YR}+\overline{T}\right)} + \frac{rc}{\alpha-r}\left(e^{(\alpha-r)\mu^{YR}}-1\right)\right] = rI. \tag{22}$$

We denote the entire term between the squared brackets by $g^{YR}\left(T^{YR}\right)$. Differentiating $f^{YR}\left(T^{YR}\right)$ with respect to $\overline{p}^{YR}$ gives:

$$\frac{\partial f^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = \frac{1}{b}g^{YR}\left(T^{YR}\right) - \frac{a-\overline{p}^{YR}}{b}\frac{\partial g^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}}, \tag{23}$$

where

$$\frac{\partial g^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = e^{-r\mu^{YR}}\left(1 - r\overline{p}^{YR}\frac{\partial \mu^{YR}}{\partial \overline{p}^{YR}} + rce^{\alpha\mu^{YR}}\frac{\partial \mu^{YR}}{\partial \overline{p}^{YR}}\right). \tag{24}$$

Recall that $\mu^R = \frac{\ln\left(\frac{\overline{p}_2^R}{c}\right)}{\alpha}$, and therefore $\frac{\partial \mu^{YR}}{\partial \overline{p}^{YR}} = \frac{1}{\alpha\overline{p}^{YR}}$. Furthermore, $e^{\alpha\mu^{YR}} = \frac{\overline{p}^{YR}}{c}$. Using this relationship we get $\frac{\partial g^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = e^{-r\mu^{YR}}$. Since $r > 0$ and $\mu^{YR} > 0$, we can infer $0 < \frac{\partial g^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} < 1$. Moreover, as

$$f^{YR}\left(T^{YR}\right) = \left(-\frac{a-\overline{p}^{YR}}{b}\right)g^{YR}\left(T^{YR}\right) = rI, \tag{25}$$

and thus $f^{YR}\left(T^{YR}\right) > 0$, where the optimality condition is fulfilled, assuming that this is possible, it follows that $g^{YR}\left(T^{YR}\right) < 0$, where the optimality condition is fulfilled, i.e. in the relevant range, given that $a > \overline{p}^{YR}$. In total, we conclude that $\frac{\partial f^{YR}(T^{YR})}{\partial \overline{p}^{YR}} < 0$. As $\frac{df^{YR}(T)}{dT} > 0$, it is obvious that $f^{YR}(T)$ is strictly monotonically increasing in T. This implies $\frac{\partial T^{YR}}{\partial \overline{p}^{YR}} > 0$ in the relevant range, where the optimality condition is fulfilled. Obviously, the participation constraint is only met, if the regulated price, $\overline{p}^{YR}$, allows full cost recovery of the investment. ■

This proposition is non-trivial and a bit surprising at first sight. It says, in words, that a lower yardstick *accelerates* the optimal replacement investment date, not the other way around. This may seem counterintuitive. If the yardstick price is lowered, the first intuition is that the discounted post-investment profits decrease. Therefore, at first glance, investment should take place later rather than earlier as a result of a lower yardstick price.

However, we can see from Propostion 2 that the story is a bit more complicated. To understand the result, two points have to be kept in mind. First, one has to look at discounted profits after investment *and* at discounted profits before investment. It is the sum of both which is of interest when looking for the optimal investment date, T^{YR}, for a given $\overline{p}^{YR}$. Second, if $\overline{p}^{YR}$ is lowered, we should compare the change of discounted profits before and after investment *at the margin*, i.e. the change of discounted profits before and after investment near $\overline{p}^{YR}$.

Indeed, a lower yardstick necessarily reduces discounted post-investment profits. Lowering the yardstick price must reduce the differences between the regulated price and marginal costs at any point in time after investment. In addition, the increase in the quantity demanded at any point in time cannot be large enough to compensate for this effect. Otherwise, the regulated firm would have lowered its price below the yardstick price before, which is impossible, since the constraint was binding before the change in the yardstick price. Moreover, due to the lower price, the cut-off point, T_S^{YR}, for yardstick regulation is reached earlier. In other words, μ^{YR} is reduced. On the other hand, an analogous logic applies to the first period. The differences between the regulated price and marginal costs also become smaller before investment. Even worse for the firm, for a strictly positive discount rate, r, the effect on discounted profits before investment, for a given change in $\overline{p}^{YR}$ and a given period in time to be considered, is even stronger than after investment due to discounting. In addition, due to the lower $\overline{p}^{YR}$, the optimal lenght of the ante-investment period is indirectly shortened anyway, since μ^{YR} is reduced, and to restore optimality, there is a negative effect on T^{YR}. The net effect of a lower $\overline{p}^{YR}$ in total is given by Proposition 2, i.e. a lower yardstick price accelerates the investment date in the relevant range.

Note the following three points. First, Proposition 2 only holds for the case of wear and tear, i.e. replacement investment. We will see below that the effect is in the opposite direction for expansion investment. Second, the yardstick level should allow cost recovery of the investment. Otherwise, the investment will not be undertaken at all. Third, using a bit of calibration we can see that lowering the yardstick price may accelerate the investment date towards the investment date maximizing discounted social welfare and, in fact, possibly to an investment date before the social welfare maximizing date, even allowing cost recovery.

Proposition 3 *For the case of wear and tear and for yardstick regulation, there may be a range, where the investment date, T_{CR}^{YR}, under yardstick regulation, with $\overline{p}^{YR}$ sufficiently high to allow cost recovery, CR, is earlier than the investment date under discounted social welfare maximization: $T_{CR}^{YR} < T^{SW}$.*

Proof. One example for which the proposition holds suffices. Assume the following parameter values: $a = 100$, $b = 1$, $c = 40$, $I = 1000$, $r = 0.07$, $\alpha = 0.01$, $\overline{T} = 10$, and $\overline{p}^{YR} = 50$. Therefore, $T^{YR} \approx 4.7$, $V_2\left(T^{YR}\right) = \int_{T^{YR}}^{T_S^{YR}} \Pi_2^{YR}(t)\, e^{-rt} dt = 2621$, $Ie^{-rT^{YR}} = 722$, which implies $V_2 - Ie^{-rT^{YR}} > 0$. Thus, $T_{CR}^{YR} = T^{YR} \approx 4.7$. With $T^{SW} \approx 7$, we get $T_{CR}^{YR} < T^{SW}$. ■

This is a result which is relevant to current policy issues. It shows that yardstick regulation can lead to inefficiently early replacement investment, even under cost recovery.

3.2.3 Replacement investment and cost-based regulation

Now, we are able to compare price-based regulation to cost-based regulation. Assume that allowed prices follow average costs. Then, strictly speaking, these prices do not necessarily have to increase after investment. With wear and tear, the decrease of variable costs per unit may actually be larger than the increase in fixed costs per unit due to investment taking into account depreciation rules, which we do not specify, and the value of time. Also, with demand growth, scale effects can lower unit costs. Yet, we consider such cases as exceptional. We concentrate on the case where allowed prices increase after investment. Thus, we assume that, under cost-based regulation, the following relationship holds: $\overline{p}_2^R > \overline{p}_1^R$.

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

$$\overline{p}^{YR} = \gamma \overline{p}_1^R + (1-\gamma)\overline{p}_2^R, \tag{26}$$

where γ 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. (26) yields:

$$\overline{p}_2^R = \frac{\overline{p}^{YR} - \gamma \overline{p}_1^R}{1-\gamma}. \tag{27}$$

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, $\overline{p}_2^R$, the above definition guarantees that the ante-investment price, $\overline{p}_1$, goes down, while the weighted average remains at $\overline{p}^{YR}$.

Proposition 4 *For the case of wear and tear, cost-based regulation accelerates the optimal investment date, T^{CB}, compared to yardstick regulation where marginal revenues are non-positive. Defining $\triangle p = \overline{p}_2^R - \overline{p}_1^R$, we can infer that $\frac{dT^{CB}}{d\triangle p} < 0$. 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:

$$\max_T V^R(T) = V_1(T) + V_2(T) - Ie^{-rT}, \tag{28}$$

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:

$$-\Omega\left(T^R\right) - \Pi_1^R\left(T^R\right) = rI, \tag{29}$$

where e^{-rT^R} was deleted. Define $z\left(T^R\right) = -\Omega\left(T^R\right) - \Pi_1^R\left(T^R\right)$. We would like to know what happens if, starting at $\overline{p}^{YR}$, we increase $\overline{p}_2^R$ and, subsequently, decrease $\overline{p}_1^R$. Thus, we examine: $\frac{\partial z(T^R)}{\partial \overline{p}_2^R} = -\frac{\partial \Omega(T^R)}{\partial \overline{p}_2^R} - \frac{\partial \Pi_1^R(T^R)}{\partial \overline{p}_2^R}$. Assuming, without loss of generality, $\gamma = 0.5$, from Eq.(26), we know that $\frac{d\overline{p}_1^R}{d\overline{p}_2^R} = -1$,which leads to:

$$\frac{\partial z\left(T^R\right)}{\partial \overline{p}_2^R} = -\frac{\partial \Omega\left(T^R\right)}{\partial \overline{p}_2^R} + \frac{\partial \Pi_1^R\left(T^R\right)}{\partial \overline{p}_1^R}. \tag{30}$$

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

$$\Omega\left(T^R\right) = \left[\overline{p}_2^R\left(e^{-r\mu^R} - 1\right) + \left(\frac{rc}{\alpha - r}\right)\left(e^{(\alpha - r)\mu^R} - 1\right)\right]\frac{a - \overline{p}_2^R}{b}. \tag{31}$$

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

$$\frac{\partial \Omega\left(T^R\right)}{\partial \overline{p}_2^R} = -\frac{1}{b(r-\alpha)}\left(\left(a - 2\overline{p}_2^R\right)(r-\alpha)\left(1 - e^{-r\mu^R}\right) + r\left(e^{r\mu^R}c - \overline{p}_2^R\right)e^{-r\mu^R}\right). \tag{32}$$

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

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

$$\frac{\partial \Pi_1^R\left(T^R\right)}{\partial \overline{p}_1^R} = \frac{a - 2\overline{p}_1^R + ce^{\alpha(T^R+\overline{T})}}{b}. \tag{33}$$

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

Substituting $\frac{\partial \Omega\left(T^R\right)}{\partial \overline{p}_2^R}$ and $\frac{\partial \Pi_1^R\left(T^R\right)}{\partial \overline{p}_1^R}$ into Eq.(30), we conclude that

$$\frac{\partial z\left(T^R\right)}{\partial \overline{p}_2^R} > 0, \tag{34}$$

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 $\frac{\partial z\left(T^R\right)}{\partial \overline{p}_2^R} > 0$, 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 $\overline{T} = 10$, and using $\overline{p}^{YR} = 60$, gives $T^{YR} \approx 7.8$. Introducing a cost-based approach, with $\overline{p}_1^{CB} = 58$ and $\overline{p}_2^{CB} = 62$, we find that $T^{CB} \approx 2.6 < T^{SW} \approx 7$. ■

The sufficiency condition to derive the result, i.e. that marginal revenues are non-positve, makes perfect sense. Proposition 4 holds at least, if marginal revenues are below zero. As we are dealing with cases of binding regulation, this is absolutely reasonable. For large values of marginal revenues, the effect reverses. If the yardstick price is at the monopoly level, a further increase of $\overline{p}_2^R$ would actually decrease post-investment profits (and the other way around for ante-investment profits). In other words, there is a level of the yardstick price beyond which an increase of $\overline{p}_2^R$ and the subsequent decrease of $\overline{p}_1^R$ 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, in contrast to 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 is similar to the case of wear and tear, but there are 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, there is actually no stopping point. Therefore, we can simplify the analysis by substituting infinity for the endpoint. For a more formal treatment, we refer to Borrmann and Brunekreeft (2009). 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, $\overline{K}$. 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 instead focuses exclusively on timing.

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:

$$\max_T V^{SW}(T) = \int_0^T SW_1(t)\, e^{-rt} dt + \int_T^\infty SW_2(t)\, e^{-rt} dt - Ie^{-rT}, \tag{35}$$

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

$$\frac{(a-c)^2}{2b} e^{gT^{SW}} - (a-c)\,\overline{K} + \frac{1}{2} b e^{-gT^{SW}} \overline{K}^2 = rI \tag{36}$$

In Eq.(36), substitute T for T^{SW} and denote the LHS by $h^{SW}(T)$.

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

$$\max_T V^{\Pi}(T) = \int_0^T \Pi_1(t)\, e^{-rt} dt + \int_T^\infty \Pi_2(t)\, e^{-rt} dt - Ie^{-rT}. \tag{37}$$

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

$$\frac{(a-c)^2}{4b} e^{gT^{\Pi}} - (a-c)\,\overline{K} + b e^{-gT^{\Pi}} \overline{K}^2 = rI. \tag{38}$$

In Eq.(38), substitute T for T^{Π} and denote the LHS by $h^{\Pi}(T)$.

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 Proposition 1 in the case of wear and tear and has been discussed extensively in Borrmann and Brunekreeft (2009).

Proposition 5 *For the case of demand growth, the optimal investment date, T^{SW}, under discounted social welfare maximization is unambiguously earlier than the optimal investment date, T^{Π}, under unregulated discounted profit-maximization: $T^{SW} < T^{\Pi}$.*

Proof. As above we examine the optimality conditions using $h^{SW}(T)$ and $h^{\Pi}(T)$ at the intersection points with rI. Both $h^{SW}(T)$ and $h^{\Pi}(T)$ are strictly convex in T and have minima. Furthermore, $\frac{dh^{SW}(T)}{dT} > 0$ at $T = 0$. Therefore, the minimum of $h^{SW}(T)$ must be at $T < 0$. Moreover, $\frac{dh^{\Pi}(T)}{dT} = 0$ at $T = 0$ and thus $\frac{dh^{\Pi}(T)}{dT} > 0$ at $T > 0$. Now, we are able to compare $h^{SW}(T)$ and $h^{\Pi}(T)$ for a given value of $\overline{K}$ and a given value of T, and we can show that $h^{SW}(T) > h^{\Pi}(T)$:

$$h^{SW}(T) - h^{\Pi}(T) = \frac{(a-c)^2}{4b} e^{gT} - \frac{1}{2} b e^{-gT} \overline{K}^2. \tag{39}$$

Substituting the extreme values first, i.e. $T = 0$ and $\overline{K} = \overline{K}_{\max} = \frac{a-c}{2b}$, gives $\frac{(a-c)^2}{8b} > 0$. The extreme value for $\overline{K}$ is the unconstrained monopoly quantity.

If $\overline{K}$ is larger than that, the outcome would be unconstrained in the unregulated case, which is in the non-relevant range. Since this also holds for $T > 0$ and $\overline{K} < \overline{K}_{\max}$, we can infer that $h^{SW}(T) > h^{\Pi}(T)$ for any $T > 0$, and thus $T^{SW} < T^{\Pi}$. ■

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.[3] Our approach to address this problem is as follows. The profit of the regulated firm is determined by the regulated prices, $\overline{p}_1^R$ and $\overline{p}_2^R$, 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, $\overline{p}_1^R < p_1$ at $Q_1 = \overline{K}$, and $\overline{p}_2^R = p_2$ 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

[3] This is a well-known problem for severely capacity-constrained airports. See, for instance, Starkie (2008).

a rule, scarce capacity is auctioned. The owners are not allowed to retain the auction revenue over and above the cost 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: $\Pi_1(t) = \left(\overline{p}_1^R - c\right) Q_1^R(t)$, with $Q_1^R(t) = \overline{K}$. It is obvious that $\left(p_1 - \overline{p}_1^R\right)\overline{K}$ is not part of the profit.

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

$$\max_T V^R(T) = \int_0^T (\overline{p}_1^R - c) Q_1^R(t) e^{-rt} dt + \int_T^\infty (\overline{p}_2^R - c) Q_2^R(t) e^{-rt} dt - I e^{-rT}. \quad (40)$$

Inserting the quantities,

$$Q_1^R(t) = \overline{K}, \text{ and } Q_2^R(t) = \frac{\left(a - \overline{p}_2^R\right) e^{gt}}{b}, \quad (41)$$

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:

$$\frac{\left(\overline{p}_2^R - c\right)\left(a - \overline{p}_2^R\right) e^{gT^R}}{b} - \left(\overline{p}_1^R - c\right)\overline{K} = rI. \quad (42)$$

In Eq.(42), substitute T for T^R and denote the LHS by $h^R(T)$.

4.2.2 Expansion investment and price-based regulation

With these preparations, we are now ready to characterize the yardstick outcome. Using the yardstick formulation, i.e. $\overline{p}^{YR} = \overline{p}_1^R = \overline{p}_2^R$, we are able to state the following proposition.

Proposition 6 *For the case of demand growth under yardstick regulation and in the relevant range, a higher regulated price accelerates the optimal investment date:* $\frac{dT^{YR}}{d\overline{p}^{YR}} < 0$. *The relevant range is from the price level equal to marginal cost up to some specific level,* p^M*, which is below the level of an unregulated monopolist maximizing discounted profits. Above the level of* p^M*, a higher yardstick price decelerates the investment date.*

Proof. We substitute $\overline{p}^{YR} = \overline{p}_1^R = \overline{p}_2^R$ in $h^R(T)$ and rewrite, which leads to the following optimality condition:

$$\left[\frac{\left(-\left(\overline{p}^{YR}\right)^2 + (a+c)\,\overline{p}^{YR} - ac\right) e^{gT^{YR}}}{b}\right] - \left(\overline{p}^{YR} - c\right)\overline{K} = rI. \quad (43)$$

This condition describes the optimal investment date, T^{YR}, maximizing discounted profits.

In Eq.(43), substitute T for T^{YR} and denote the LHS by $h^{YR}(T)$.

Similarly, substitute T for T^{YR} in the expression in squared brackets and denote the term by $\omega^{YR}(T)$. Bear in mind that the relationship $\omega^{YR}(T) > 0$ must hold to allow the optimality condition to be fulfilled and that $\frac{d\omega^{YR}(T)}{dT} > 0$ for $c < \overline{p}^{YR} < \frac{a+c}{2}$. This implies that $h^{YR}(T)$ is strictly monotonically increasing in T in this range. Note that $h^{YR}(T) = 0$ for $\overline{p}^{YR} = c$.

In addition, focussing on the optimal investment date, T^{YR}, we can derive the following results:

$$\frac{\partial h^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = \frac{\left(-2\overline{p}^{YR} + a + c\right) e^{gT^{YR}}}{b} - \overline{K}. \tag{44}$$

Using the extreme values, i.e. $\overline{p}^{YR} = c$, $T^{YR} = 0$ and $\overline{K} = \frac{a-c}{2b}$, we find:

$$\frac{\partial h^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = \frac{(a-c)\, e^{gT^{YR}}}{b} - \overline{K} \geq 0. \tag{45}$$

This relationship also holds for $T^{YR} \geq 0$ and $\overline{K} \leq \frac{a-c}{2b}$. Furthermore, for $\overline{p}^{YR} = \frac{a+c}{2}$, i.e. at the price which an unregulated monopolist maximizing discounted profits sets, $\frac{\partial h^{YR}\left(T^{YR}\right)}{\partial \overline{p}^{YR}} = -\overline{K} < 0$. Moreover,

$$\frac{\partial^2 h^{YR}\left(T^{YR}\right)}{\left(\partial \overline{p}^{YR}\right)^2} = \frac{-2e^{gT^{YR}}}{b} < 0. \tag{46}$$

Thus, we can see that $h^{YR}\left(T^{YR}\right)$ is strictly concave in $\overline{p}^{YR}$ at T^{YR}, with a maximum at some level of $\overline{p}^{YR}$ between c and $\frac{a-c}{2b}$. ■

Starting at a reasonably low price level, i.e. starting at $\overline{p}^{YR} = c$, the optimal investment date is accelerated when $\overline{p}^{YR}$ rises. This effect vanishes at some level, p^M, which is below the price level of an unregulated monopolist maximizing discounted profits. Then, the effect is reversed, i.e. the investment date is decelerated with a further increase of $\overline{p}^{YR}$. We call the interval $c \leq \overline{p}^{YR} \leq p^M$ the relevant range. Summing up, we find that $\frac{dT^{YR}}{d\overline{p}^{YR}} < 0$ for the relevant range of $\overline{p}^{YR}$. The relevant range applies, if we take regulation seriously, since the relevant range is close to cost-oriented prices, and it only ends near the unregulated monopoly price. Thus, Proposition 6 states that, for the relevant cases of binding regulation aiming at just and reasonable prices, a higher allowed price and therefore a higher rate of return on investment, accelerates expansion investment. It may be noted that this result is the mirror opposite of the wear and tear case for replacement investment, where a higher yardstick decelerates the optimal investment date.

Proposition 7 *For the case of demand growth under yardstick regulation, the effect of yardstick regulation on the optimal investment date, as compared to unregulated monopoly, is ambiguous:* $T^{YR} \lesseqgtr T^{\Pi}$.

Proof. A numerical example which shows that both possibilities exist suffices. Use the following parameter values: $a = 100$, $b = 1$, $c = 40$, $I = 25000$, $r = 0.07$, $g = 0.05$, $\overline{T} = 10$, and $\overline{K} = 20$. These parameter values yield $p^{\Pi} = 70$ and $T^{\Pi} \approx 22.9$. Now, assume $\overline{p}^{YR} = 50$, which implies $T^{YR} \approx 27.2$, and therefore $T^{YR} > T^{\Pi}$. Alternatively, assume $\overline{p}^{YR} = 60$, which leads to $T^{YR} \approx 19.8$, and thus $T^{YR} < T^{\Pi}$. ■

The first part of the Proof of Proposition 7, i.e. $T^{YR} > T^{\Pi}$, can be demonstrated more elegantly, if we compare $h^{\Pi}(T)$ to $h^{YR}(T)$ for $\overline{p}^{YR} = c$ for a given value of T. Since $h^{YR}(T) = 0 < h^{\Pi}(T)$ at $\overline{p}^{YR} = c$, we can infer that $T^{YR} > T^{\Pi}$. In fact, perhaps more telling, for $\overline{p}^{YR} = c$, note that, as $h^{YR} \longrightarrow 0$, $T^{YR} \longrightarrow \infty$, or, in words, investment would simply not take place, because the regulated price would be too low. This means that too low prices will decelerate expansion investment so much that, effectively, it will not take place at all.

The second part of the proof of Proposition 7, i.e. $T^{YR} < T^{\Pi}$, is the more surprising part. It implies that, even for expansion investment, yardstick regulation can accelerate investment as compared to the unregulated case. However, we emphasize that this holds for relatively high prices only and that it is the exception rather than the rule. The basic intuition with expansion investment is that the regulated private monopoly invests later than the unregulated monopoly, which in turn invests later than is socially optimal.

However, there are details which modify the basic intuition, and as suggested by Proposition 7, the optimal investment date for a regulated monopoly can be earlier than for an unregulated monopoly. First, assume $\overline{K} = 0$. This case is straightforward. As the regulated yardstick approaches the unregulated monopoly price, we find, unambiguously, that the investment dates converge. There is essentially no difference. Therefore, for $\overline{K} = 0$, $T^{YR} \geq T^{\Pi}$. Things change, if $\overline{K} > 0$, in which case the unregulated investor eats away part of her own ante-investment profit. Assume a capacity-constrained unregulated monopoly making an expansion investment. By mechanism of the high constrained price before investment, the post-investment profit-maximizing price will be lower than the ante-investment price. This also affects the revenues of the capacity which is already there (at $\overline{K}$). The unregulated monopoly will take these lower revenues on existing capacity into account, and compensating this means to delay the investment date. For the yardstick-regulated monopoly, this reasoning does not apply. If, for comparison, we assume that the yardstick level is at the level of the unregulated post-investment profit-maximizing price, then, by definition of yardstick regulation, this must also be the ante-investment price. Therefore, the ante-investment level of the yardstick would be lower than the unregulated constrained capacity level and therefore the yardstick regulated firm would have less to lose on current capacity, $\overline{K}$. Thus, we conclude that, since the yardstick-regulated investor has less to lose on existing assets, yardstick regulation may actually accelerate the investment date as compared to the investment date for the unregulated investor.

However, with reasonably low cost-oriented regulated prices, it is highly likely that yardstick regulation decelerates the investment date compared to the case of the absence of regulation. Only for unreasonably high regulated prices, we find that yardstick regulation can accelerate the investment date compared to the case of the absence of regulation. There are two opposing effects at work. First, for $\overline{p}^{YR} < p_2^{\Pi}$, if $\overline{p}^{YR}$ is raised, Π^{YR} increases and therefore T^{YR} goes down. Since $\Pi^{YR} < \Pi^{\Pi}$, we nevertheless conclude that $T^{YR} > T^{\Pi}$. This is the basic effect, which says that, if the regulated price goes up, the investment moment is accelerated. Second, as explained above, the unregulated monopolist will decrease the post-investment price as compared to the capacity-constrained ante-investment price: $p_2^{\Pi} < p_1^{\Pi}$. This implies that T^{YR} can be earlier than T^{Π}. The second effect can dominate the first effect, if $\overline{p}^{YR}$ is high, since then the first effect is small. However, as long as we take regulation seriously, $\overline{p}^{YR}$ will not be sufficiently high, and the first effect is likely to dominate. Therefore, we take $T^{YR} > T^{\Pi}$ as the normal case and $T^{YR} < T^{\Pi}$ as the exceptional case.

Moreover, as we will explore in somewhat more detail below, the second effect completely vanishes, if $\overline{K} = 0$.

Conjecture: *For the case of demand growth under yardstick regulation, $T^{YR} > T^{SW}$. In words, we conjecture that, under yardstick regulation, the timing of expansion investment is always decelerated as compared to the timing of socially optimal expansion investment .*

We restrict ourselves to a conjecture, which seems fairly plausible, though. In the Proof of Proposition 6, there is a specific value, p^M, of the yardstick price, $\overline{p}^{YR}$, beyond which investment is decelerated. However, without specifying p^M, which we did not do, it is difficult to prove the conjecture. Specifying p^M turns out to be tedious and unclear. More straightforward is the following approximation, which comes close to a proof. We use extreme values, which brings the investment date close to the earliest investment date under yardstick regulation and show that, for these values, the conjecture holds. Take the capacity level, $\frac{a-c}{2b}$, of an unregulated monopolist maximizing discounted profits. Also, use the price level $\frac{a+c}{2}$, i.e. $\overline{p}^{YR} = \frac{a+c}{2}$.

Substitute these values into $h^{SW}(T)$ and $h^{YR}(T)$ and check, for each given T, whether $h^{SW}(T) > h^{YR}(T)$, implying that $T^{SW} < T^{YR}$. This yields:

$$\frac{(a-c)^2}{2b}e^{gT} - \frac{(a-c)^2}{4b} + \frac{(a-c)^2}{8b}e^{-gT} > \frac{(a-c)^2}{4b}e^{gT}. \tag{47}$$

This gives:

$$2e^{gT} - 2 + e^{-gT} > 0, \tag{48}$$

which always holds for $T > 0$. Therefore, it is very unlikely that yardstick regulation induces inefficiently fast expansion investment. Quite the contrary, if anything, we have to be concerned that yardstick regulation inefficiently decelerates expansion investment. Thus, we now turn to cost-based regulation.

4.2.3 Expansion investment and cost-based regulation

Using the mechanism to compare yarstick-regulated with cost-based regulation as defined in Section 3.2.3, with $\overline{p}^{YR} = \gamma \overline{p}_1^R + (1-\gamma)\overline{p}_2^R$ and $\triangle p = \overline{p}_2^R - \overline{p}_1^R$, we are able to state the following proposition.

Proposition 8 *For the case of demand growth, assuming $\overline{p}_2^R > \overline{p}_1^R$, cost-based regulation accelerates the investment date for $\overline{K} > 0$ compared to price-based regulation, whereas the investment dates for cost-based regulation and price-based regulation are equal for $\overline{K} = 0$. The following relationship holds: $\frac{dT^{CB}}{d\triangle p} < 0$.*

Proof. This result is similar to the first part of Proposition 4. Building on the formulation of the objective function in Eq.(40), we find immediately that

$$\Pi_2^R\left(T^{CB}\right) - \Pi_1^R\left(T^{CB}\right) = rI. \tag{49}$$

Defining $Z\left(T^{CB}\right) = \Pi_2^R\left(T^{CB}\right) - \Pi_1^R\left(T^{CB}\right)$, it is obvious that $\frac{\partial Z\left(T^{CB}\right)}{\partial \overline{p}_2^R} = \frac{\partial \Pi_2^R\left(T^{CB}\right)}{\partial \overline{p}_2^R} - \frac{\partial \Pi_1^R\left(T^{CB}\right)}{\partial \overline{p}_1^R}\frac{\partial \overline{p}_1^R}{\partial \overline{p}_2}$. As in the Proof of Proposition 4, we assume, without loss of generality, that $\frac{\partial \overline{p}_1^R}{\partial \overline{p}_2^R} = -1$. Since $\frac{\partial \Pi_2^R\left(T^{CB}\right)}{\partial \overline{p}_2^R} > 0$ and $\frac{\partial \Pi_1^R\left(T^{CB}\right)}{\partial p_1^R} > 0$, it can be easily seen that $\frac{\partial \Pi_1^R\left(T^{CB}\right)}{\partial \overline{p}_2^R} < 0$. Thus, for a given value of rI, we get $\frac{\partial Z\left(T^{CB}\right)}{\partial \overline{p}_2^R} > 0$, 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 $T^{CB} \lesseqgtr T^{SW}$. 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 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: $a = 100$, $b = 1$, $c = 20$, $I = 25,000$, $r = 0.07$, $g = 0.05$, $\overline{T} = 10$, and $\overline{K} = 30$. Use $\overline{p}_1^R = 20$, and $\overline{p}_2^R = 60$. This gives $T^{SW} \approx 3.27$ and $T^{CB} \approx 1.79$, and therefore $T^{CB} < T^{SW}$. 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. $\overline{K} > 0$.

Proposition 9 *For the case of demand growth and for $\overline{K} = 0$, we find $\frac{dT^{CB}}{d\overline{p}_2^R} < 0$ and $T^{CB} (= T^{YR}) \geq T^{\Pi} > T^{SW}$.*

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 $\overline{K} = 0$.

The case of $\overline{K} = 0$ 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 $\overline{K} = 0$. 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 top-ups 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 $\overline{K} = 0$. The propositions above suggest that this policy will indeed accelerate expansion investment.

For the case of $\overline{K} = 0$, the effects on timing are unambiguous, as claimed in Proposition 9. As the acceleration effect of Proposition 7 now vanishes, there is only a deceleration effect as compared to the unregulated case. Therefore, we conclude that, for expansion investment with $\overline{K} = 0$, 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, which brings us back to the discussion in Borrmann and Brunekreeft (2009).

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. The concerns are threefold. 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. In fact, there are many examples of investment needs at present.

The issue of efficient investment incentives is at the core of the current debate on regulation. Whereas it receives a lot of attention in practice, the theoretical literature is, up to now, rather silent. Our paper contributes to the theoretical literature by exploring the relationship between regulation and monopoly investment. We examine the differences between price-based regulation and cost-based regulation. Within the literature on the relationship between regulation and investment, we focus on the *timing* of investment. More specifically, we study the investment issue under *certainty* and concentrate on *monopoly* investment, which, in our case, means large investment outlays of a fixed nature, so that timing is an issue at all. Our main conclusions are the following.

First, the optimal investment date of an unregulated monopoly maximizing discounted profits, given high investment outlays which are independent of output, is decelerated compared to the socially optimal investment date. This result holds equally for replacement investment and for expansion investment. It was discussed extensively in Borrmann and Brunekreeft (2009). Indeed, the structure of the problem is reminiscent of the well-known classical problem of a static situation with (cost) economies of scale, where welfare-optimal prices equal to marginal costs imply that cost recovery is impossible. In this case, unregulated private provision of a good without subsidies cannot lead to first-best prices. Instead, second-best prices with full cost recovery are an obvious alternative from a normative point of view. The problem discussed in this paper is essentially a dynamic variation of this.

Second, we draw several conclusions on the effects of pure price-based regulation on the timing of replacement investment and the timing of expansion investment. By pure price-based regulation, we mean yardstick regulation, where allowed prices or revenues are not related to the underlying own costs of the regulated firm, but rather depend on benchmarking with something external to the firm, following Shleifer (1985). Thus, an investment cannot be passed through into higher prices. We note that this seems counterintuitive for the monopoly domain, but normal in a competitive setting. Your typical cornershop cannot increase its prices just because it invests, since customers can simply go to another cornershop. This, of course, reminds us of what yardstick regulation aims to do: mimic competition. However, it may go wrong, as we are dealing with monopoly investment with high investment oulays of a fixed nature. A detailed analysis gives unambiguous insights. On the one hand, yardstick regulation may work just fine for replacement investment, if properly applied. In fact, we show that a lower yardstick price accelerates the investment date, provided that the price level still allows full cost recovery. On the other hand, yardstick regulation does not work well for expansion investment. A lower yardstick price may decelerate expansion investment inefficiently.

Third, the aim of our paper is to examine, whether cost-based regulation can accelerate the optimal investment date compared to price-based regulation. The answer to this question is unambiguous. Cost-based regulation can accelerate the optimal investment date compared to price-based regulation, but this may not always be efficient from a social point of view. In other words, especially for replacement investment, a cost-based approach can quite easily accelerate the privately optimal investment date inefficiently fast. In still 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. Thus, acceleration triggered by a cost-based approach to regulation may improve efficiency. 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. Cost-based regulation, including rate-of-return adders and top-ups, may speed up investment. However, as far as the efficiency of investment timing is important, details matter.

Two related assumptions are crucial to our model and suggest topics for further research. Our approach assumes a monopoly situation, and it neither presumes a race for investment nor tendering. 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.

Furthermore, this paper relies on high investment outlays, which are independent of output, and constant marginal costs at any point in time. As can be seen from the different optimality conditions in different settings, this assumption drives many results. First, the differences in timing become small, if investment outlays become small. Second, the assumption of high investment outlays which are independent of output justifies quite naturally our assumption that investment takes place only once. Third, if investment outlays are relatively low, an adequate model should allow repeated investment. We expect that, if investment outlays are stepwise decreased and if cost functions are strictly convex, optimization results in the normal competitive outcome, and timing differences vanish. We leave this for further reseach, but note that the investment timing problem, as worked out in this paper, is genuinely a problem of monopoly investment with high investment outlays, which are independent of output.

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] Borrmann, J. and Brunekreeft, G. (2009), The timing of monopoly investment, mimeo, University of Vienna and Jacobs University Bremen.

[5] Brunekreeft, G. (2004), Market-based investment in electricity transmission networks: Controllable flow, *Utilities Policy*, **12** (4), 269-281.

[6] Brunekreeft, G. (2005), Regulatory issues in merchant transmission investment, *Utilities Policy*, **13** (2), 175-186.

[7] Brunekreeft, G. (2009), Regulation of network charges, in *Handbook Utility Management*, Bausch, A. and Schwenker, B., eds., Springer: Berlin, 435-456.

[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] Joskow, P.L. (2006), Incentive regulation in theory and practice: Electricity distribution and transmission networks, Cambridge Working Paper Series in Economics, 0607, University of Cambridge.

[18] Katz, M.L. and Shapiro, C. (1987), R&D rivalry with licensing or imitation, *American Economic Review*, **77** (3), 402-420.

[19] Nagel, T. and Rammerstorfer, M. (2009), Modeling investment behavior under price cap regulation, *Central European Journal of Operations Research,* **17** (2), 111-129.

[20] Shleifer, A. (1985), A theory of yardstick competition, *RAND Journal of Economics*, **16** (3), 319-327.

[21] Spence, A.M. (1975), Monopoly, quality, and regulation, *Bell Journal of Economics*, **6** (2), 417-429.

[22] Starkie, D. (2008), *Aviation Markets: Studies in Competition and Regulatory Reform*, Ashgate: Aldershot, Hampshire.

[23] Vickers J. and Yarrow, G. (1988), *Privatization: An Economic Analysis*, MIT Press: Cambridge, Mass.

[24] Vogelsang, I. (2009), Incentive regulation, investments and technological change, mimeo, Boston University.