Bremer Energie Institut

Bremen Energy Working Papers

Paperseries No. 08

Jörg Borrmann and Gert Brunekreeft

The Timing of Repeated and Unrepeated Monopoly Investment under Wear and Tear and Demand Growth

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

Ansprechpartner:

Nele Friedrichsen

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

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

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

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

The Timing of Repeated and Unrepeated Monopoly Investment under Wear and Tear and Demand Growth*

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

Gert Brunekreeft
Jacobs University Bremen
College Ring 2
28759 Bremen
Germany

May 4, 2011

Abstract

In an intertemporal model, we analyze the timing of irreversible and lumpy monopoly investment under certainty. There are two reasons for investing, i.e. wear and tear leading to replacement investment and demand growth leading to expansion investment. Both in a single investment setting and in a repeated investment setting, we find that a firm maximizing discounted social welfare invests earlier than an identical firm maximizing discounted profits. The investment date of an identical firm maximizing a discounted convex combination of social welfare and profits lies between these polar cases. All results apply both to replacement investment and to expansion investment.

Keywords: Expansion investment, Investment timing, Monopoly, Repeated investment, Replacement investment.

JEL classification: D42, G00, L20

*The authors gratefully acknowledge useful comments on earlier versions of the paper by participants at the "Centre for Competition and Regulatory Policy Workshop" at the University of Birmingham in July 2007, at the 2008 Business & Economics Society International Conference in Lugano in July 2008, at the 36th Annual Conference of the European Association for Research in Industrial Economics at the University of Ljubljana in September 2009, at the Workshop "Current Topics in the Regulation of Energy and Telecommunications Markets" at Vienna University of Economics and Business in January 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

In spite of the liberalization process which has taken place in many infrastructure sectors since the early 1980s, e.g. energy, telecommunications, and transportation, there are many examples of networks which are still monopolized. For instance, electricity and gas transmission networks, electricity and gas distribution networks, road networks, and infrastructure networks in the rail sector involving tracks, bridges, tunnels, train stations and terminals are, in general, still monopolies.

From a practical perspective, the timing of monopoly investment is definitely relevant to current economic policy issues. If a monopoly firm invests earlier than it should, society as a whole incurs opportunity costs due to resources allocated suboptimally. In case that the firm invests later than it should, this may have detrimental consequences due to foregone benefits. The debate on the deteriorating quality of energy networks, e.g. in the United States, is a case in point.

From a theoretical perspective, the timing of monopoly investment is also crucial. Using conventional static analysis, it can be shown that a monopolist maximizing profits chooses inefficiently high prices compared to a monopolist maximizing social welfare, as is well-known to any reader of textbooks explaining the fundamentals of economics (see, e.g., Samuelson and Nordhaus, 2010). Thus, the question arises, if there is an intertemporal counterpart of this static result. In other words, it is to be found out, whether an intertemporal model can show that a monopoly maximizing discounted profits invests inefficiently late or inefficiently early compared to a monopoly maximizing discounted social welfare, or not.

In this paper, we add to the literature by examining the timing of monopoly investment under wear and tear and under demand growth in a single framework. In an intertemporal model, we analyze the behavior of a monopoly firm in different situations. We address the cases of a firm maximizing discounted social welfare, a firm maximizing discounted profits, and a firm maximizing a discounted convex combination of social welfare and profits. In our model, there are two reasons for investing. The first reason is wear and tear leading to replacement investment. The second reason is demand growth leading to expansion investment. Investment is both lumpy and completely irreversible in the sense that there is no alternative use for the assets after investing. We assume certainty.

In our basic scenario, a monopoly firm aims to invest in productive assets only once. Investing necessitates an initial outlay, at a single point in time. Both for wear and tear and for demand growth, we find that a firm maximizing discounted social welfare invests earlier than an identical firm maximizing discounted profits. The investment date of an identical firm maximizing a discounted convex combination of social welfare and profits lies between these polar cases.

In order to address the case of repeated investment, we generalize our basic scenario. Starting from an initial optimal investment date we derive the next optimal investment date. As this procedure can be applied infinitely, we are thus able to describe the case of infinitely repeated investment. We find that the time interval between two successive investments of a monopoly firm maximizing discounted social welfare is shorter than the corresponding time interval of an identical firm maximizing discounted profits. Analogous to our basic scenario, the corresponding time interval between two successive investments of an identical firm maximizing a discounted convex combination of social welfare and profits lies between these two polar cases.

Strictly speaking, we analyze the effects of the *objective function* of a monopoly firm on the timing of investment. However, our results can also be understood in a broader sense. First, one can interpret a monopoly firm which maximizes discounted social welfare as a (perfect) publicly-owned monopolist and a monopoly firm which maximizes discounted profits as an (unregulated) privately-owned monopolist. Correspondingly, one can interpret a monopoly firm which maximizes a discounted convex combination of social welfare and profits as a monopolist with mixed ownership, with the weighting factor of the convex combination reflecting the ownership shares. Mixed ownership can often be found in network industries, for instance in the telecommunications sector in Austria and in the energy sector in Germany. To the best of our knowledge, a suitable framework for the analysis of the effects of mixed ownership on investment timing does not exist up to now. All in all, this is an interpretation of our model in terms of *ownership*. Second, as is well-known from the literature on optimal regulation, there are surplus subsidy schemes designed for privately-owned monopolists assumed to maximize (discounted or undiscounted) total profits, i.e. the (discounted or undiscounted) sum of their operating profits and some sort of subsidy. These schemes serve the purpose of molding privately motivated actions into socially desirable outcomes. In particular, Loeb and Magat's (1979) mechanism deserves to be mentioned, which results in the first-best outcome by maximizing total surplus in a partial equilibrium model. Their static mechanism can also be used to achieve intertemporal objectives. This is an interpretation of our model in terms of *regulation*. Third, private sector monopoly firms can, by differentiating prices, lower the difference between profits and total surplus. This is an interpretation of our model in terms of *price discrimination*.

The paper is organized as follows. In Section 2, we give a brief overview of the related literature. Afterwards, we introduce our basic scenario model with a firm aiming to invest only once in Section 3. In Section 4, we analyze replacement investment, and in Section 5, we analyze expansion investment. We extend our analysis to repeated investment in Section 6. In Section 7, we conclude.

2 Related literature

An early treatment of the problem of investment timing is Marglin (1963), who suggests that, under certain conditions, an investment with a negative net present value can be transformed into one with a positive net present value by means of delay.

Another early contribution is Terborgh (1949), who develops an approximate decision rule for the optimal replacement of assets which deteriorate taking into account technological progress. He assumes that technological progress is embodied in new equipment and examines when it is optimal to replace the oldest capital with the most recent vintages. Smith (1961) elaborates on Terborgh's (1949) analysis. Malcomson (1975) also analyzes the optimal replacement policy of a firm. In particular, he asks if the optimal machine life is given uniquely by cost conditions, whether it is independent of output and product market conditions or not, and if it is constant over time, as is asserted by Terborgh (1949) and Smith (1961). Malcomson's (1975) results are extended by van Hilten (1991). Whereas Malcomson (1975) provides numerical evidence that the optimal machine life is constant under certain conditions, van Hilten (1991) makes these conditions more explicit and gives an analytical proof of the constancy of the optimal machine life. Tarola and Trento (2010) analyze the plant-size problem by allowing firms to combine their investment policy with a pricing policy. Assuming that a firm's equipment is affected by physical deterioration which reduces production capacity over time and embedding technological advancements in their model, Tarola and Trento (2010) derive several general properties of the optimal pricing policy and the ensuing optimal sequence of investments of the firm over time.

Another long-standing issue in the investment literature is the plant-size problem, i.e. the question of how to add facilities to meet rising demand in industries characterized by economies of scale in investment costs, which is addressed by Chenery (1952), Manne (1961, 1967) and Srinivasan (1967). Manne (1967) assumes an exogenous demand growing linearly over time and finds that the cost-minimal investment policy is characterized by the so-called constant-cycle property, which means that successive investments are all of the same size and undertaken at equally spaced points in time. Srinivasan (1967) shows that the constant-cycle property also holds for a situation where demand grows geometrically over time. Several refinements and many applications to public services are developed later (see, e.g., Luss, 1982; Li and Tirupati, 1992; Nam and Logendran, 1992). In addition, a considerable amount of research focuses on the problems of lead times and uncertain demand (see, e.g., Nickell, 1977; Freidenfelds, 1981, Bean et al., 1992, Chaouch and Buzacott, 1994; Ryan, 2004). Demichelis and Tarola (2006) extend the framework for the study of the plant-size problem by allowing the monopoly firm to combine its investment policy with a pricing policy adjusting demand upwards or downwards over time. They derive several properties of the optimal investment and pricing policies under the assumptions that the firm maximizes discounted profits and that it, while controlling the size of the investments to be undertaken, does not control the investment dates which are assumed to be equally spaced on the time axis. Alternatively, one can assume that the monopoly firm may decide on the dates, but not on the investment sizes. Tarola (2006) shows that, in this case, the optimal pricing policy leads the firm to manipulate the price pattern in order to ensure that the dates at which the fixed investments have to be undertaken are equally spaced over time.

Timing decisions are also examined in the literature on innovation and patent races. Important contributions are, for instance, Dasgupta and Stiglitz (1980), Reinganum (1981), Katz and Shapiro (1987), and Fudenberg and Tirole (1987). An excellent survey is given by Reinganum (1989). A similar kind of analysis is used by Szymanski (1991) to explore the timing of infrastructure investment.

Finally, the timing of monopoly investment is also analyzed under the assumption that the firm is regulated. One branch of the literature going in this direction is Dobbs (2004) and, following up on that, Nagel and Rammerstorfer (2009), where investment takes place under uncertainty. Their models rely on the literature on real options. Another branch follows from a debate on the concept of access holidays (see, in particular, Gans and Williams, 1999; Gans and King, 2004).

Our paper is related both to the literature on the optimal replacement policy and to the literature on the plant-size problem. However, unlike both branches of the traditional literature, we address the two main drivers for investment in a single framework, i.e. wear and tear on the one hand, and demand growth on the other hand. Furthermore, we analyze the interaction between a monopoly firm's objectives, its pricing policy, and its investment policy. Like Tarola (2006), we assume that the firm may make decisions on the investment dates but not on the investment sizes. However, unlike her, we distinguish explicitly between wear and tear and demand growth, whereas she focuses on demand growth. Furthermore, she confines her analysis to a monopolist maximizing discounted profits, while we compare the effects of different objectives of the firm on investment timing. Our analysis is also related to Tarola and Trento (2010), who analyze the replacement problem by deriving several general properties of the optimal pricing policy and the ensuing optimal sequence of investments of a firm over time. While, in their model, physical deterioration reduces production capacity over time, we analyze the replacement problem by assuming that marginal cost rises over time. Furthermore, unlike Tarola and Trento (2010), we derive explicit solutions for the optimal timing problem, both for the case of a single investment and for the case of repeated investment. Our intertemporal model also builds on the innovation literature, in particular Katz and Shapiro (1987), as well as on Szymanski (1991), on Gans and Williams (1999), and, in particular, on Brunekreeft and Newbery (2006). However, we also analyze repeated investment, whereas the latter three focus on a single investment project. Unlike large parts of the regulation literature, we do not take into account uncertainty. We do this on purpose in order to be able to isolate possible effects of different objectives of the monopoly firm on investment timing. Thus, we are able to identify the intertemporal counterpart of static market failure for a monopoly firm.

3 Basic scenario model

For simplicity, we consider a *single*-product monopolist. In our view, this assumption is satisfactory for our purposes, i.e. to adequately examine the interaction between the objective of the firm, its pricing policy over time and its investment policy over time. However, our results can be easily generalized to a multi-product monopolist.

In our basic scenario model, which we develop in this section and apply in the following two sections, the monopoly firm aims to invest in productive assets merely *once*. This assumption can be justified by the fact that investment assets in monopolized industries are typically characterized by high investment outlays and long lives. For instance, electricity distribution networks can be assumed to operate for 40 years. Nevertheless, this is merely the first step of our analysis. In the second step, we extend our analysis to the case of repeated investment in Section 6.

To keep the model as simple as possible, we assume that investing necessitates merely an initial outlay, I, with $I \in \mathbb{R}^+$, *at a single point in time.*

Furthermore, we assume the initial investment outlay, I, to be *fixed*. It is no decision variable of the firm. This assumption is a simple way to capture the feature that investment in monopolized infrastructure sectors is typically lumpy. If I were perfectly divisible, there would be nothing special about monopoly investment. The solution to the problem of investment timing would be trivial. It is the assumption of lumpy investment which drives our results.

In our model, investment is assumed to be *completely irreversible* in the sense that there is no alternative use for the assets after investing.

We assume that the *discount rate*, denoted by r, with $r \in \mathbb{R}^+$, is *independent of the objective functions* of decision makers. This implies that the discount rate which is relevant to private investors is identical to the social discount rate which is relevant to a social planner. The independence of the discount rate is an assumption, which is neither self-evident nor uncontroversial. We choose this assumption to make the results of our analysis comparable.

The objective function of the firm is either *discounted social welfare*, or *discounted profits*, or a *discounted convex combination of social welfare and profits*. We define social welfare as total surplus, i.e. as the sum of consumer surplus and profits. This ensures that marginal cost prices maximize social welfare, under rather general conditions, in a partial equilibrium model. We are aware of the fact that there are situations where this social welfare measure does not seem appropriate, in particular, when income effects are relevant (see, e.g., Tirole, 1988; Armstrong et al., 1994). However, given that the social welfare measure of total surplus is useful due to its simplicity and that the conditions under which it is inadequate can be precisely described (see, in particular, Willig, 1976), we prefer it over alternatives.

For the ease of exposition, we assume that the *participation constraints* are *always met*. The investment allows the firm to attain strictly positive discounted social welfare, strictly positive discounted profits, or a strictly positive discounted convex combination of social welfare and profits, respectively.

In our basic scenario model, 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. In our intertemporal model, there are two reasons for investing. The first reason is *wear and tear* leading to replacement investment. The second reason is *demand growth* leading to expansion investment.

3.1 Wear and tear

Production costs, $C(\cdot,\cdot,\cdot)$, not taking into account I, are a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of time, t. We assume that, at any point in time, marginal costs are constant. The age of the existing assets at $t = 0$ is denoted by $\overline{T}$. Over time, either marginal costs do 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_I, brings marginal cost back to its original level:

$$C(Q_1, Q_2, t) = \begin{cases} ce^{\alpha\left(t+\overline{T}\right)}Q_1; & t < T_I \\ ce^{\alpha(t-T_I)}Q_2; & T_I \leq t, \end{cases} \tag{1}$$

where $c \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$. By defining costs in this way, we ensure that there are economies of scale due to investment at any point in time. Furthermore, we assume $r > \alpha$ to ensure stability.

3.2 Demand growth

Inverse demand, $P(\cdot,\cdot)$, is a function of the quantity demanded, Q, and of time, t. For simplicity, we assume that, at any point in time, demand is *linear*.

At first sight, a possible alternative assumption could be that demand is constant elasticity. Technically, this would have the advantage that, with rising price levels, the quantity demanded would never drop to zero. However, due to the cost function we choose, the assumption of a constant demand elasticity would not fit our model. It would imply that there is no profit-maximizing price. Furthermore, a strictly concave demand function is no convincing alternative, either. With a strictly concave demand function, we would also run into unnecessary problems as a consequence of our definition of demand growth, which we will explain in turn.

Over time, either the relationship between the quantity demanded and price does not change, i.e. $g = 0$, or the quantity demanded at a given price grows at a constant rate, g, with $0 < g < 1$:

$$P(Q,t) = a - be^{-gt}Q, \tag{2}$$

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

This is a particular pattern of demand growth. The demand curve rotates around the point where it intersects the vertical axis. A potential alternative is, obviously, to assume parallel shifts of the demand curve over time. That assumption is, for instance, common in the investment literature on the plant-size problem. We prefer to use our specification because of the following reason. Parallel shifts of linear demand curves due to a change in a non-price variable (e.g., income) are not compatible with the assumption of a constant elasticity with respect to that variable (e.g., a constant income elasticity). If the elasticity with respect to some non-price variable shall be constant, then a demand curve which is linear must rotate when the non-price variable changes (see, for instance, Graves and Sexton, 2006). To ensure stability, we assume $r > g$.

3.3 Methodology

In order to separate the effects of wear and tear and the effects of demand growth, we consider two different versions of our basic scenario model. In the first version, marginal costs increase at a constant rate, $0 < \alpha < 1$, *and* the relationship between the quantity demanded and price does not change over time, i.e. $g = 0$. Thus, in the first version, there is only a reason for replacement investment, and no reason for expansion investment. In the second version, marginal costs do not change over time, i.e. $\alpha = 0$, *and* the quantity demanded at a given price grows at a constant rate, g, with $0 < g < 1$. This implies that, in the second version, there is only a reason for expansion investment, and no reason for replacement investment.

4 Wear and tear without demand growth: replacement investment

4.1 A monopoly firm maximizing discounted social welfare

Assertions of the optimality of marginal cost prices have a long history in economics (e.g., Spulber, 1989). For instance, in a partial equilibrium framework, marginal cost prices are, under appropriate conditions, optimal in the sense that they maximize the social welfare measure we choose, i.e. total surplus (or, in other words, the sum of aggregate consumer surplus and profits), on a particular market. Therefore, one might argue that marginal cost prices should be the aim of a social planner from an efficiency perspective, at least if there are no distortions elsewhere in the economy, or if the particular sector in question is sufficiently separated from the rest of the economy (e.g., Waterson, 1988).

However, there are also some drawbacks of the concept of marginal cost prices like measurement problems, inefficiencies, and conceptual difficulties. In particular, if there are cost economies of scale, as is the case with our cost function, government subsidies need to be available in order to allow the firm to break even, at least in the absence of price discrimination. This raises the difficult question of the shadow cost of public funds (e.g., Laffont and Tirole, 1993).

Nevertheless, given that the rationale behind marginal prices can easily be understood and that marginal cost prices are firmly based on cost, it is not by coincidence that they are the traditional benchmark for public utility pricing (e.g., Mitchell and Vogelsang, 1991). For instance, one might think of a publicly-owned monopoly that is subsidized and acts in the best interest of society. Alternatively, one might have in mind a privately-owned monopoly under Loeb and Magat's (1979) surplus subsidy scheme, or a privately-owned monopoly using primary price discrimination (also called perfect price discrimination). In this section, we follow the tradition of total surplus and add an intertemporal component, thus aiming to maximize discounted total surplus.

Aggregate consumer surplus, $CS(\cdot,\cdot)$, at any point in time is a function of the quantity demanded, Q, and of time, t. With (2) and $g = 0$, it can be written as:

$$CS(Q,t) = \frac{b}{2}Q^2, \tag{3}$$

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

Profits, $\Pi(\cdot,\cdot,\cdot)$, at any point in time, not taking into account I, are a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of time, t. With (1) and (2), as well as $g = 0$, we obtain:

$$\Pi\left(Q_1, Q_2, t\right) = \begin{cases} -bQ_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1; & t < T_I \\ -bQ_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2; & T_I \leq t, \end{cases} \tag{4}$$

where $a, b, c \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

Total surplus, $TS\left(\cdot,\cdot,\cdot\right)$, at any point in time, not taking into account I, is the sum of (3) and (4):

$$TS\left(Q_1, Q_2, t\right) = \begin{cases} -\frac{b}{2}Q_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1; & t < T_I \\ -\frac{b}{2}Q_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2; & T_I \leq t, \end{cases} \tag{5}$$

where $a, b, c \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

A monopolist aiming to maximize discounted total surplus will not use his assets infinitely, since marginal cost rises over time, whereas inverse demand does not change. When a becomes smaller than marginal cost, $ce^{\alpha(t-T_I)}$, after investment, the firm will no longer produce. It is straightforward to show that it will stop production at $T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}$.

Building on (5), discounted total surplus, $DTS\left(\cdot,\cdot,\cdot\right)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of the investment date, T_I:

$$\begin{aligned} DTS\left(Q_1, Q_2, T_I\right) = & \int_0^{T_I} \left(-\frac{b}{2}Q_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1\right) e^{-rt} dt \\ & + \int_{T_I}^{T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}} \left(-\frac{b}{2}Q_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2\right) e^{-rt} dt \\ & - Ie^{-rT_I}, \end{aligned} \tag{6}$$

where $a, b, c, r \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

4.2 A monopoly firm maximizing discounted profits

Now, we describe a monopoly firm which maximizes discounted profits. Its profit function is given by (4). Just like a firm which maximizes discounted social welfare, a monopoly firm maximizing discounted profits will not use its assets infinitely. It will also stop production at $T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}$.

Building on (4), discounted total profits, $D\Pi\left(\cdot,\cdot,\cdot\right)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of the investment date, T_I:

$$D\Pi\left(Q_1, Q_2, T_I\right) = \int\limits_0^{T_I} \left(-bQ_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1\right) e^{-rt} dt \qquad (7)$$

$$+ \int\limits_{T_I}^{T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}} \left(-bQ_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2\right) e^{-rt} dt$$

$$-Ie^{-rT_I},$$

where $a, b, c, r \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

4.3 A monopoly firm maximizing a discounted convex combination of social welfare and profits

In the following, we generalize the approaches of Section 4.1 and Section 4.2 by addressing a monopoly firm which maximizes a discounted convex combination of social welfare and profits. Total surplus is weighted by γ, and profits are weighted by $(1-\gamma)$; $0 \leq \gamma \leq 1$. The special case of a firm maximizing discounted total surplus is given by $\gamma = 1$, the special case of a firm maximizing discounted profits is given by $\gamma = 0$, and all the cases in between are characterized by $0 < \gamma < 1$.

A possible interpretation of the case where $0 < \gamma < 1$ is that of a *mixed monopoly*, which is a monopoly that is partly publicly-owned and partly privately-owned and that acts in the best interest of both categories of shareholders. In this interpretation of the model, the weighting factors, γ and $(1-\gamma)$, of the convex combination, reflect the shares of the public and the private owners, respectively. As already mentioned, mixed ownership can often be found in network industries, for instance, if former public firms have been partially privatized.

Building on (4) and (5), the mixed monopoly objective, $MM\left(\cdot,\cdot,\cdot\right)$, at any point in time (not taking into account I) is a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of time, t:

$$MM\left(Q_1, Q_2, t\right) = \gamma TS\left(Q_1, Q_2, t\right) + \left(1-\gamma\right)\Pi\left(Q_1, Q_2, t\right), \qquad (8)$$

where $Q_1, Q_2, t \in \mathbb{R}_0^+$. This is equivalent to:

$$MM\left(Q_1, Q_2, t\right) = \begin{cases} \left(\frac{\gamma}{2} - 1\right) bQ_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1; & t < T_I \\ \left(\frac{\gamma}{2} - 1\right) bQ_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2; & T_I \leq t, \end{cases} \qquad (9)$$

where $a, b, c \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

A firm which maximizes a discounted convex combination of social welfare and profits will also stop production at $T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}$.

Building on (9), the discounted mixed monopoly objective, $DMM(\cdot,\cdot,\cdot)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_1, before investment, of the outputs, Q_2, after investment, and of the investment date, T_I:

$$\begin{aligned} DMM(Q_1, Q_2, T_I) = & \int_0^{T_I} \left(\left(\frac{\gamma}{2}-1\right) bQ_1^2 + \left(a - ce^{\alpha\left(t+\overline{T}\right)}\right) Q_1\right) e^{-rt}dt \\ & + \int_{T_I}^{T_I+\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} \left(\left(\frac{\gamma}{2}-1\right) bQ_2^2 + \left(a - ce^{\alpha(t-T_I)}\right) Q_2\right) e^{-rt}dt \\ & -Ie^{-rT_I}, \end{aligned} \tag{10}$$

where $a, b, c, r \in \mathbb{R}^+$ and $Q_1, Q_2, \overline{T}, T_I, t \in \mathbb{R}_0^+$.

4.4 A comparison of the investment dates

With (10) we can analyze the behavior of a firm which maximizes a discounted convex combination of total surplus and profits.

Proposition 1 *The time paths of outputs, Q_1^{DMM} and Q_2^{DMM}, which maximize a discounted convex combination of total surplus and profits for a given weighting factor, γ, result from the corresponding convex combination of marginal cost prices and profit-maximizing prices at any point in time.*

Proof. See Appendix. ■

Proposition 2 *The investment date, T_I^{DMM}, which maximizes a discounted convex combination of total surplus and profits for a given weighting factor, γ, is given by:*

$$T_I^{DMM} = \frac{1}{\alpha} \ln \left(\frac{a - \sqrt[2]{a^2 + \frac{2acr}{\alpha - r} - \frac{c^2 r}{2\alpha - r} - \frac{\alpha^2 2a^2}{\left(r^2 - 3\alpha r + 2\alpha^2\right)\left(\frac{a}{c}\right)^{\frac{r}{\alpha}}} - (4-2\gamma)\, brI}}{ce^{\alpha\overline{T}}} \right). \tag{11}$$

Proof. See Appendix. ■

Using $\gamma = 1$ to describe the investment date, T_I^{DTS}, of a firm maximizing discounted total surplus, and $\gamma = 0$ to describe the investment date, $T_I^{D\Pi}$, of an identical firm maximizing discounted profits, it is easy to derive the following fundamental result.

Corollary 3 *In terms of wear and tear and other things being equal, the investment date, T_I^{DTS}, which maximizes discounted total surplus is earlier than the investment date, T_I^{DMM}, which maximizes a discounted convex combination of total surplus and profits for a given weighting factor, γ, with $0 < \gamma < 1$, and is earlier than the investment date, $T_I^{D\Pi}$, which maximizes discounted profits.*

5 Demand growth without wear and tear: expansion investment

Now, we turn to demand growth as a driver of expansion investment. In this scenario, wear and tear is absent, whereas the quantity demanded at a given price grows at a constant rate.

At the point of departure, there is an *old capacity limit*, K_1, before expansion investment takes place. We assume that the old capacity limit, K_1, is *exogenously given*. We choose this assumption to simplify matters, as we do not want to focus on optimal capacity choice. The assumption of an exogenously given capacity limit also reflects our view that monopoly investment is typically lumpy. The old capacity limit, K_1, is assumed to have just become binding. Therefore, the firm will invest in order to expand capacity to a *new capacity limit*, K_2. Analogous to the old capacity limit, K_1, the new capacity limit, K_2, is also *exogenously given*. At the *investment date*, T_I, the new capacity limit, K_2 is assumed to be *greater than the quantity demanded*. However, due to growing demand, the new capacity limit, K_2, is assumed to become *binding later on*. It is obvious that the monopoly firm will *use its assets infinitely*, since marginal cost does not rise over time.

5.1 A monopoly firm maximizing discounted total surplus

We denote the points in time, when the capacity limits, K_1 and K_2, become binding constraints by $T_1^{bind,DTS}$ and $T_2^{bind,DTS}$. With (2) and $\alpha = 0$, aggregate consumer surplus, $CS(\cdot,\cdot)$, at any point in time (not taking into account the initial investment outlay, I) can easily be determined. It is a function of the outputs, Q_2, immediately after investment, and of time, t. We obtain:

$$CS(Q_2, t) = \begin{cases} \frac{b}{2} e^{-gt} K_1^2; & T_1^{bind,DTS} \leq t < T_I \\ \frac{b}{2} e^{-gt} Q_2^2; & T_I \leq t < T_2^{bind,DTS} \\ \frac{b}{2} e^{-gt} K_2^2; & T_2^{bind,DTS} \leq t, \end{cases} \tag{12}$$

where $a, b, c \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,DTS}, T_2^{bind,DTS}, T_I, t \in \mathbb{R}_0^+$.

Profits, $\Pi(\cdot,\cdot)$, at any point in time (not taking into account I) can also be written as a function of the outputs, Q_2, immediately after investment, and of time, t. With (1) and (2) as well as $\alpha = 0$, we get:

$$\Pi\left(Q_2, t\right) = \begin{cases} -be^{-gt}K_1^2 + (a-c)\,K_1; & T_1^{bind,DTS} \leq t < T_I \\ -be^{-gt}Q_2^2 + (a-c)\,Q_2; & T_I \leq t < T_2^{bind,DTS} \\ -be^{-gt}K_2^2 + (a-c)\,K_2; & T_2^{bind,DTS} \leq t, \end{cases} \tag{13}$$

where $a, b, c \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,DTS}, T_2^{bind,DTS}, T_I, t \in \mathbb{R}_0^+$.

Total surplus, $TS\left(\cdot,\cdot\right)$, at any point in time (not taking into account I) is the sum of (12) and (13):

$$TS\left(Q_2, t\right) = \begin{cases} -\frac{b}{2}e^{-gt}K_1^2 + (a-c)\,K_1; & T_1^{bind,DTS} \leq t < T_I \\ -\frac{b}{2}e^{-gt}Q_2^2 + (a-c)\,Q_2; & T_I \leq t < T_2^{bind,DTS} \\ -\frac{b}{2}e^{-gt}K_2^2 + (a-c)\,K_2; & T_2^{bind,DTS} \leq t, \end{cases} \tag{14}$$

where $a, b, c \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,DTS}, T_2^{bind,DTS}, T_I, t \in \mathbb{R}_0^+$.

Building on (14), discounted total surplus, $DTS\left(\cdot,\cdot\right)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_2, after investment, and of the investment date, T_I:

$$\begin{aligned} DTS\left(Q_2, T_I\right) = & \int\limits_{T_1^{bind,DTS}}^{T_I} \left(-\tfrac{b}{2}e^{-gt}K_1^2 + (a-c)\,K_1\right)e^{-rt}dt \\ & + \int\limits_{T_I}^{T_2^{bind,DTS}} \left(-\tfrac{b}{2}e^{-gt}Q_2^2 + (a-c)\,Q_2\right)e^{-rt}dt \\ & + \int\limits_{T_2^{bind,DTS}}^{\infty} \left(-\tfrac{b}{2}e^{-gt}K_2^2 + (a-c)\,K_2\right)e^{-rt}dt \\ & -Ie^{-rT_I}, \end{aligned} \tag{15}$$

where $a, b, c, r \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,DTS}, T_2^{bind,DTS}, T_I, t \in \mathbb{R}_0^+$.

5.2 A monopoly firm maximizing discounted profits

In the following, we analyze the behavior of a monopoly firm which maximizes discounted profits. We denote the points in time, when the capacity limits, K_1 and K_2, become binding constraints by $T_1^{bind,D\Pi}$ and $T_2^{bind,D\Pi}$.

Analogous to (13), profits, $\Pi\left(\cdot,\cdot\right)$, at any point in time (not taking into account I) can be written as a function of the outputs, Q_2, immediately after investment, and of time, t. Building on (1) and (2), taking into account that $\alpha = 0$, and inserting $T_1^{bind,D\Pi}$ and $T_2^{bind,D\Pi}$, we get:

$$\Pi\left(Q_2, t\right) = \begin{cases} -be^{-gt}K_1^2 + (a-c)\,K_1; & T_1^{bind,D\Pi} \leq t < T_I \\ -be^{-gt}Q_2^2 + (a-c)\,Q_2; & T_I \leq t < T_2^{bind,D\Pi} \\ -be^{-gt}K_2^2 + (a-c)\,K_2; & T_2^{bind,D\Pi} \leq t, \end{cases} \tag{16}$$

where $a, b, c \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,D\Pi}, T_2^{bind,D\Pi}, T_I, t \in \mathbb{R}_0^+$.

Building on (16), discounted total profits, $D\Pi(\cdot,\cdot)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_2, after investment, and of the investment date, T_I:

$$\begin{aligned} D\Pi(Q_2, T_I) = & \int\limits_{T_1^{bind,D\Pi}}^{T_I} \left(-be^{-gt}K_1^2 + (a-c)K_1\right)e^{-rt}dt \\ & + \int\limits_{T_I}^{T_2^{bind,D\Pi}} \left(-be^{-gt}Q_2^2 + (a-c)Q_2\right)e^{-rt}dt \\ & + \int\limits_{T_2^{bind,D\Pi}}^{\infty} \left(-be^{-gt}K_2^2 + (a-c)K_2\right)e^{-rt}dt \\ & - Ie^{-rT_I}, \end{aligned} \tag{17}$$

where $a, b, c, r \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,D\Pi}, T_2^{bind,D\Pi}, T_I, t \in \mathbb{R}_0^+$.

5.3 A monopoly firm maximizing a discounted convex combination of social welfare and profits

Now, we address a monopoly firm which maximizes a discounted convex combination of social welfare and profits. Total surplus is weighted by γ, and profits are weighted by $(1-\gamma)$; $0 \leq \gamma \leq 1$. We denote the points in time, when the (exogenously given) capacity limits, K_1 and K_2, become binding constraints by $T_1^{bind,DMM}$ and $T_2^{bind,DMM}$.

The mixed monopoly objective, $MM(\cdot,\cdot)$, at any point in time (not taking into account I) is a function of the of the outputs, Q_2, after investment, and of time, t:

$$MM(Q_2, t) = \gamma TS(Q_2, t) + (1-\gamma)\Pi(Q_2, t), \tag{18}$$

where $Q_2, t \in \mathbb{R}_0^+$. This is equivalent to:

$$MM(Q_2, t) = \begin{cases} \left(\frac{\gamma}{2} - 1\right)be^{-gt}K_1^2 + (a-c)K_1; & T_1^{bind,DMM} \leq t < T_I \\ \left(\frac{\gamma}{2} - 1\right)be^{-gt}Q_2^2 + (a-c)Q_2; & T_I \leq t < T_2^{bind,DMM} \\ \left(\frac{\gamma}{2} - 1\right)be^{-gt}K_2^2 + (a-c)K_2; & T_2^{bind,DMM} \leq t, \end{cases} \tag{19}$$

where $a, b, c \in \mathbb{R}^+$ and $K_1, K_2, Q_2, T_1^{bind,DMM}, T_2^{bind,DMM}, T_I, t \in \mathbb{R}_0^+$.

Building on (19), the discounted mixed monopoly objective, $DMM(\cdot,\cdot)$, over time, taking into account the initial investment outlay, I, can be written as a function of the outputs, Q_2, after investment, and of the investment date, T_I:

$$
\begin{aligned}
DMM\left(Q_2,T_I\right) = & \int\limits_{T_1^{bind,DMM}}^{T_I} \left(\left(\tfrac{\gamma}{2}-1\right) be^{-gt}K_1^2+(a-c)\,K_1\right) e^{-rt}dt \\
& + \int\limits_{T_I}^{T_2^{bind,DMM}} \left(\left(\tfrac{\gamma}{2}-1\right) be^{-gt}Q_2^2+(a-c)\,Q_2\right) e^{-rt}dt \\
& + \int\limits_{T_2^{bind,DMM}}^{\infty} \left(\left(\tfrac{\gamma}{2}-1\right) be^{-gt}K_2^2+(a-c)\,K_2\right) e^{-rt}dt \\
& -Ie^{-rT_I},
\end{aligned}
\tag{20}
$$

where $a,b,c,r \in \mathbb{R}^+$ and $K_1,K_2,Q_2,T_1^{bind,DMM},T_2^{bind,DMM},T_I,t \in \mathbb{R}_0^+$.

5.4 A comparison of the investment dates

Referring to (20) we can analyze the behavior of a firm which maximizes a discounted convex combination of total surplus and profits.

Proposition 4 *The time path of outputs, Q_2^{DMM}, immediately after investment which maximize a discounted convex combination of total surplus and profits for a given weighting factor, γ, results from the corresponding convex combination of marginal cost prices and profit-maximizing prices at any point in time.*

Proof. See Appendix. ■

Proposition 5 *The investment date, T_I^{DMM}, which maximizes a discounted convex combination of total surplus and profits for a given weighting factor, γ, is given by:*

$$
T_I^{DMM} = \frac{\ln\left((2-\gamma)\,b\frac{rI+(a-c)K_1+\sqrt[2]{(rI+(a-c)K_1)^2-(a-c)^2K_1^2}}{(a-c)^2}\right)}{g}. \tag{21}
$$

Proof. See Appendix. ■

Using $\gamma=1$ to describe the investment date, T_I^{DTS}, of a firm maximizing discounted total surplus, and $\gamma=0$ to describe the investment date, $T_I^{D\Pi}$, of an identical firm maximizing discounted profits, it is easy to derive the following fundamental result.

Corollary 6 *In terms of demand growth and other things being equal, the investment date, T_I^{DTS}, which maximizes discounted total surplus is earlier than the investment date, T_I^{DMM}, which maximizes a discounted convex combination of total surplus and profits for a given weighting factor, γ, with $0 < \gamma < 1$, and is earlier than the investment date, $T_I^{D\Pi}$, which maximizes discounted profits.*

Corollary 6 is analogous to Corollary 3, which covers the case of wear and tear.

6 Repeated investment model

In this section, we will generalize our basic scenario model to repeated investment. Starting from an initial optimal investment date we derive the next optimal investment date. As this procedure can be applied infinitely, we are thus able to describe the case of infinitely repeated investment.

Now, we assume that *every* investment necessitates an outlay, I, with $I \in \mathbb{R}^+$, *at a single point in time*. Again, this investment outlay, I, is assumed to be *fixed*. Furthermore, we assume that I does not change over time, which means that this amount is invested at each investment date.

In our repeated investment model, the firm has to make several decisions simultaneously. It has to decide *which outputs to set* after investment, and *when to invest* in the assets. Contrary to the basic scenario model, the firm does not need to decide how long to use the assets, since the time interval between two investments is always shorter than the production period.

6.1 Wear and tear

Let T_n be an investment date, which is the starting point of our analysis. Then, denote the future investment dates by T_{n+1}, T_{n+2}, T_{n+3}, etc. At each investment date, the fixed amount to be invested is I, the investment outlay. Let Q_{n+1} be the output produced between T_n and T_{n+1}, let Q_{n+2} be the output produced between T_{n+1} and T_{n+2}, let Q_{n+3} be the output produced between T_{n+2} and T_{n+3}, etc. All other variables are defined just like in the basic scenario model.

Then, analogous to (10), discounted total surplus can be characterized as:

$$
\begin{aligned}
& DMM\left(Q_{n+1}, Q_{n+2}, Q_{n+3}, \ldots, T_n, T_{n+1}, T_{n+2}, T_{n+3}, \ldots\right) \\
& = \int\limits_{T_n}^{T_{n+1}} \left(\left(\tfrac{\gamma}{2}-1\right) bQ_{n+1}^2 + \left(a - ce^{\alpha(t-T_n)}\right) Q_{n+1}\right) e^{-rt} dt - Ie^{-rT_{n+1}} \\
& + \int\limits_{T_{n+1}}^{T_{n+2}} \left(\left(\tfrac{\gamma}{2}-1\right) bQ_{n+2}^2 + \left(a - ce^{\alpha(t-T_{n+1})}\right) Q_{n+2}\right) e^{-rt} dt - Ie^{-rT_{n+2}} \quad , \\
& + \int\limits_{T_{n+2}}^{T_{n+3}} \left(\left(\tfrac{\gamma}{2}-1\right) bQ_{n+3}^2 + \left(a - ce^{\alpha(t-T_{n+2})}\right) Q_{n+3}\right) e^{-rt} dt - Ie^{-rT_{n+3}} \\
& \vdots
\end{aligned} \tag{22}
$$

where $0 < \alpha < 1$ and $0 \leq \gamma \leq 1$ and $a, b, c, r \in \mathbb{R}^+$ and $Q_{n+1}, Q_{n+2}, Q_{n+3}, \ldots, T_n, T_{n+1}, T_{n+2}, T_{n+3}, \ldots, t \in \mathbb{R}_0^+$.

We are looking for the optimal investment dates, i.e. for those investment dates, $T_{n+1}, T_{n+2}, T_{n+3}, \ldots$, which maximize discounted total surplus. Obviously, the first-order conditions can be described as:

$$
\frac{\partial DMM}{\partial T_{n+1}} = \frac{\partial DMM}{\partial T_{n+2}} = \frac{\partial DMM}{\partial T_{n+3}} = \ldots = 0. \tag{23}
$$

However, to simplify matters, we confine ourselves to the analysis of T_{n+1} assuming that T_n is an optimal investment date, i.e. an investment date, T_n^{DMM}, maximizing a discounted convex combination of total surplus and profits for a given weighting factor, γ. In other words, assuming $T_n = T_n^{DMM}$ we aim to characterize the next optimal investment date, i.e. T_{n+1}^{DMM}.

This characterization makes it possible to compare identical monopoly firms with different objectives.

Proposition 7 *In terms of wear and tear and other things being equal, the time interval between two successive investments is smaller for a monopolist maximizing discounted total surplus than for an identical monopolist maximizing a discounted convex combination of total surplus and profits for a given weighting factor, γ, with $0 < \gamma < 1$, for whom it is smaller than for an identical monopolist maximizing discounted profits.*

Proof. See Appendix. ■

6.2 Demand growth

Let T_n be the *initial investment date*, which is the starting point of our analysis. As a consequence of this initial investment, the firm expands some exogenously given *initial capacity limit*, K_n, to the exogenously given *second capacity limit*, K_{n+1}. At the outset, K_{n+1} is greater than the quantities demanded. However, as demand grows, the second capacity limit, K_{n+1}, becomes binding at some point in time. Therefore, the monopoly firm later expands capacity at a specific point in time, i.e. at the *second investment date*, T_{n+1}, to the exogenously given *third capacity limit*, K_{n+2}. This capacity limit is assumed to be greater than the quantities demanded in the beginning. However, as demand grows, the third capacity limit becomes binding at some point in time. Therefore, the monopoly firm later expands capacity at a specific point in time, i.e. at the *third investment date*, T_{n+2}, to the exogenously given *fourth capacity limit*, K_{n+3}, etc. For simplicity, we assume that $K_{n+1} - K_n = K_{n+2} - K_{n+1} = K_{n+3} - K_{n+2} = \ldots = \Delta K$. At each investment date, the fixed amount to be invested is I, the investment outlay. Furthermore, we denote the quantities demanded between the initial investment date, T_n, and the second investment date, T_{n+1}, by Q_n, the quantities demanded between T_{n+1} and T_{n+2} by Q_{n+1}, the quantities demanded between T_{n+2} and T_{n+3} by Q_{n+2}, etc. Obviously, $Q_n = K_{n+1}$, after the second capacity limit, K_{n+1}, has become binding, $Q_{n+1} = K_{n+2}$, after the third capacity limit, K_{n+2}, has become binding, $Q_{n+2} = K_{n+3}$, after the fourth capacity limit, K_{n+3}, has become binding, etc. We denote the points in time, when the exogenously given capacity limits, K_{n+1}, K_{n+2}, K_{n+3}, etc. become binding constraints by $T_n^{bind,DMM}$, $T_{n+1}^{bind,DMM}$, $T_{n+2}^{bind,DMM}$, etc., respectively. All other variables are defined just like in the basic scenario model.

Then, analogous to (20), a discounted convex combination of total surplus and profits for a given weighting factor, γ, can be characterized as:

$$
\begin{aligned}
&DMM\left(Q_n, Q_{n+1}, Q_{n+2}, Q_{n+3}, \ldots, T_n, T_{n+1}, T_{n+2}, T_{n+3}, \ldots\right) \\
&= \int\limits_{T_n}^{T_n^{bind,DMM}} \left(\left(\tfrac{\gamma}{2} - 1\right) be^{-gt} Q_n^2 + (a-c)\, Q_n\right) e^{-rt} dt \\
&+ \int\limits_{T_n^{bind,DMM}}^{T_{n+1}} \left(\left(\tfrac{\gamma}{2} - 1\right) be^{-gt} K_{n+1}^2 + (a-c)\, K_{n+1}\right) e^{-rt} dt - Ie^{-rT_{n+1}} \\
&+ \int\limits_{T_{n+1}}^{T_{n+1}^{bind,DMM}} \left(\left(\tfrac{\gamma}{2} - 1\right) be^{-gt} Q_{n+1}^2 + (a-c)\, Q_{n+1}\right) e^{-rt} dt \\
&+ \int\limits_{T_{n+1}^{bind,DMM}}^{T_{n+2}} \left(\left(\tfrac{\gamma}{2} - 1\right) be^{-gt} K_{n+2}^2 + (a-c)\, K_{n+2}\right) e^{-rt} dt - Ie^{-rT_{n+2}} \\
&\vdots
\end{aligned}
\quad , \quad (24)
$$

where $a,\ b,\ c,\ r\ \in \mathbb{R}^+$ and $K_{n+1},\ K_{n+2},\ldots,\ Q_n,\ Q_{n+1},\ldots,T_n^{bind,DMM}$, $T_{n+1}^{bind,DMM},\ldots,T_n,T_{n+1},T_{n+2},\ldots,t\in\mathbb{R}_0^+$.

We are looking for the optimal investment dates, i.e. for those investment dates, $T_{n+1},T_{n+2},T_{n+3},\ldots$, which maximize (24). Again, we confine ourselves, however, to the analysis of T_{n+1} assuming that T_n is an optimal investment date, T_n^{DMM}. In other words, assuming $T_n=T_n^{DMM}$ we aim to describe the next optimal investment date, i.e. T_{n+1}^{DMM}.

This characterization makes it possible to compare identical monopoly firms with different objectives. It is, however, possible that the time interval between two successive investments for a monopoly firm maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ, is not constant over time. Define $x_{n+1,n}^{DMM}=T_{n+1}^{DMM}-T_n^{DMM}$. Then, potentially, $x_{n+1,n}^{DMM}\neq x_{n+2,n+1}^{DMM}\neq x_{n+3,n+2}^{DMM}$... Therefore, a meaningful comparison of identical monopoly firms with different objectives may only be possible for time intervals between *corresponding* successive investments, i.e. only for *time intervals with identical subscripts.* Let the time interval between two successive investments for an identical monopoly firm maximizing discounted total surplus corresponding to $x_{n+1,n}^{DMM}$ be $x_{n+1,n}^{DTS}=T_{n+1}^{DTS}-T_n^{DTS}$, and let the corresponding time interval between two successive investments for an identical monopoly firm maximizing discounted profits be $x_{n+1,n}^{D\Pi}=T_{n+1}^{D\Pi}-T_n^{D\Pi}$. Then, we can compare $x_{n+1,n}^{DTS}$ to $x_{n+1,n}^{DMM}$ and to $x_{n+1,n}^{D\Pi}$.

Proposition 8 *In terms of demand growth and other things being equal, the time interval between two corresponding successive investments is smaller for a monopolist maximizing discounted total surplus than for an identical monopolist maximizing a discounted convex combination of total surplus and profits for a given weighting factor, γ, with $0<\gamma<1$, for whom it is smaller than for an identical monopolist maximizing discounted profits.*

Proof. See Appendix. ■

7 Conclusion

The debate on the deteriorating quality of monopolized infrastructure networks, e.g. in the energy sector and in the rail sector in many countries throughout the world, demonstrates that the timing of monopoly investment is definitely relevant to current economic policy issues.

In practice, there are many factors influencing the timing of monopoly investment. For instance, for publicly-owned monopolies, it may be relevant, whether the government runs a budget deficit or a surplus. Furthermore, the objectives of the legislature may be ordered or diffuse. For privately-owned monopolists, it may be important, whether they are subject to investment obligations or not.

However, with few exceptions, the timing of monopoly investment has not been one of the major topics on the research agenda of public utility economics until recently. In particular, there are only very few articles on the interaction between monopoly pricing and investment timing. Given the importance of infrastructure investment for the welfare of society as a whole, our aim has been to fill the gap.

In this paper, we have addressed typical properties of monopoly investment, i.e. lumpiness and irreversibility. Furthermore, differentiating between wear and tear on the one hand and demand growth on the other hand, we have focused on replacement investment and on expansion investment. The paper delivers an analysis of both types of investment within a unified framework. Moreover, we have addressed the case of a single investment and the case of infinitely repeated investment.

Our results can be summarized as follows. For a single investment, we have been able to show that a monopolist maximizing discounted social welfare invests earlier than an identical monopolist maximizing discounted profits. The investment date of an identical firm maximizing a discounted convex combination of social welfare and profits lies between these polar cases. This applies both to wear and tear leading to replacement investment, and to demand growth leading to expansion investment. For infinitely repeated investment, we have been able to show that the time interval between two corresponding successive investments is smaller in the case of a monopolist maximizing discounted total surplus than in the case of an identical monopolist maximizing discounted profits. The corresponding time interval between two successive investments of an identical firm maximizing a discounted convex combination of social welfare and profits lies between these two polar cases. Once again, this applies both to replacement investment and to expansion investment.

How can these results be interpreted from a theoretical point of view? We have been able to identify an intertemporal counterpart of static market failure for a monopoly firm. Whereas, in a static setting, a monopolist maximizing profits chooses inefficiently high prices compared to a monopolist maximizing social welfare, in an intertemporal setting, a monopolist maximizing discounted profits invests inefficiently late compared to a monopolist maximizing discounted social welfare. The reasons underlying the monopoly behavior are similar. In a static setting, a monopoly maximizing profits does, in the absence of price discrimination, not take into account consumer surplus when deciding on prices. In an intertemporal setting, a monopoly maximizing discounted profits does not take into account discounted consumer surplus when deciding on investment dates.

Furthermore, it is important to note that the magnitudes of the differences in investment dates between a monopolist maximizing discounted social welfare and an identical monopolist maximizing discounted profits depend on the size of the investment outlay. If the investment outlay is small, the same principles may hold as in a situation where the investment outlay is high. However, for a small size of the investment outlay, we would not consider investment timing to be an exciting issue. Essentially, it is the lumpiness of investment which creates an interesting problem.

From an economic policy perspective, the trade-offs in an intertemporal setting are reminiscent of those in a static setting. Whereas a publicly-owned firm which acts in the best interest of society can, theoretically, set marginal cost prices which maximize social welfare in a partial equilibrium model and can, in principle, invest at investment dates maximizing discounted social welfare, it has to be subsidized to be able to do so. Thus, the problem of the shadow cost of public funds comes into play, if there are cost economies of scale. Furthermore, publicly-owned monopolists cannot necessarily be expected to achieve productive efficiency. A privately-owned monopolist does not require government subsidies and can be expected to produce at lower cost than a public sector monopolist. However, if unregulated, prices are inefficiently high and investment dates are inefficiently late. Mixed ownership strikes the balance between these two types of monopolists.

Thus, in the absence of price discrimination which allows the monopolist to capture (at least part of) the consumer surplus the question of price regulation arises. Frequently, it is argued that static welfare losses induced by monopoly under cost economies of scale due to inefficient pricing can, at least in principle, be lowered by price regulation, if information problems are not too severe. Thus, the question occurs whether price regulation can also be an adequate answer to the intertemporal problem of inefficiently late investment. We conclude by stating that an important area for future research will be the analysis of the relationship between monopoly price regulation and investment timing, and we would like to add that a first step into this direction is our recently completed companion paper (Brunekreeft and Borrmann, 2011).

References

[1] Armstrong, M., Cowan, S., and Vickers, J., *Regulatory Reform: Economic Analysis and British Experience*, Cambridge, Mass.: MIT Press.

[2] Bean, J.C., Higle, J., and Smith, R.L. (1992), Capacity Expansion under Stochastic Demand, *Operations Research*, **40**, 210-216.

[3] Brunekreeft, G., and Borrmann, J. (2011), The Effect of Monopoly Regulation on the Timing of Investment, *Bremen Energy Working Papers*, No. 9, May, Jacobs University Bremen.

[4] Brunekreeft, G., and Newbery, D. (2006), Should Merchant Transmission Be Subject to a Must-offer Provision?, *Journal of Regulatory Economics*, **30**, 233-260.

[5] Chaouch, B.A., and Buzacott, J.A. (1994), The Effects of Lead Time on Plant Timing and Size, *Production and Operations Management*, **3**, 38-54.

[6] Chenery, H.B. (1952), Overcapacity and the Acceleration Principle, *Econometrica*, **20**, 1-28.

[7] Dasgupta, P., and Stiglitz, J. (1980), Uncertainty, Industrial Structure, and the Speed of R&D, *Bell Journal of Economics*, **11**, 1-28.

[8] Demichelis, S., and Tarola, O. (2006), Capacity Expansion and Dynamic Monopoly Pricing, *Research in Economics*, **60**, 169-178.

[9] Dobbs, I.M. (2004), Intertemporal Price Cap Regulation under Uncertainty, *Economic Journal*, **114**, 421-440.

[10] Freidenfelds, J. (1981), *Capacity Expansion: Analysis of Simple Models with Applications*, New York: North-Holland.

[11] Fudenberg, D., and Tirole, J. (1987), Understanding Rent Dissipation: On the Use of Game Theory in Industrial Organization, *American Economic Review, P&P*, **77**, 176-138.

[12] Gans, J.S., and King, S.P. (2004), Access Holidays and the Timing of Infrastructure Investment, *Economic Record*, **80**, 89-100.

[13] Gans, J.S., and Williams, P.L. (1999), Access Regulation and the Timing of Infrastructure Investment, *Economic Record*, **75**, 127-37.

[14] Graves, P.E., and Sexton, R.L. (2006), Demand and Supply Curves: Rotations *versus* Shifts, *Atlantic Economic Journal*, **34**, 361-364.

[15] Katz, M.L., and Shapiro, C. (1987), R&D Rivalry with Licensing or Imitation, *American Economic Review*, **77**, 402-420.

[16] Laffont, J.-J., and Tirole, J. (1993), *A Theory of Incentives in Procurement and Regulation*, Cambridge, Mass.: MIT Press.

[17] Li, S., and Tirupati, D. (1992), Technology Selection and Capacity Expansion for Manufacturing Systems, in Kusiak, A., ed., *Intelligent Design and Manufacturing*, New York: John Wiley & Sons, 257-288.

[18] Loeb, M., and Magat, W.A. (1979), A Decentralized Method for Utility Regulation, *Journal of Law and Economics*, **22**, 399-404.

[19] Luss, H. (1982), Operations Research and Capacity Expansion Problems: A Survey, *Operations Research*, **30**, 907-947.

[20] Malcomson, J.M. (1975), Replacement and the Rental Value of Capital Equipment Subject to Obsolescence, *Journal of Economic Theory*, **10**, 24-41.

[21] Manne, A.S. (1961), Capacity Expansion and Probabilistic Growth, *Econometrica*, **29**, 632-649.

[22] Manne, A.S. (1967), Two Producing Areas - Constant Cycle Time Policies, in Manne, A.S., ed., *Investments for Capacity Expansion: Size, Location and Time Phasing*, Cambridge, Mass.: MIT Press, 193-209.

[23] Marglin, S.A. (1963), *Approaches to Dynamic Investment Planning*, Amsterdam: North-Holland.

[24] Mitchell, B.M., and Vogelsang, I. (1991), *Telecommunications Pricing: Theory and Practice*, Cambridge, Engl.: Cambridge University Press.

[25] Nagel, T., and Rammerstorfer, M. (2009), Modeling Investment Behavior under Price Cap Regulation, *Central European Journal of Operations Research,* **17**, 111-129.

[26] Nam, S., and Logendran, R. (1992), Aggregate Production Planning: A Survey of Models and Methodologies, *European Journal of Operational Research*, **61**, 255-272.

[27] Nickell, S. (1977), Uncertainty and Lags in the Investment Decisions of Firms, *Review of Economic Studies*, **44**, 249-263.

[28] Reinganum, J.F. (1981), Market Structure and the Diffusion of New Technology, *Bell Journal of Economics*, **12**, 618-624.

[29] Reinganum, J.F. (1989), The Timing of Innovation: Research, Development, and Diffusion, in Schmalensee, R., and Willig, R.D., eds., *Handbook of Industrial Organization*, Vol. I, Amsterdam: North-Holland, 849-908.

[30] Ryan, S.M. (2004), Capacity Expansion for Random Exponential Demand Growth with Lead Times, *Management Science*, **50**, 740-748.

[31] Samuelson, P.A., and Nordhaus, W.D. (2010), *Economics*, 19th ed., New York: McGraw-Hill.

[32] Smith, V.L. (1961), *Investment and Production: A Study in the Theory of the Capital-Using Enterprise*, London: Oxford University Press.

[33] Spulber, D.F. (1989), *Regulation and Markets*, Cambridge, Mass.: MIT Press.

[34] Srinivasan, T.N. (1967), Geometric Rate of Growth of Demand, in Manne, A.S., ed., *Investments for Capacity Expansion: Size, Location and Time Phasing*, Cambridge, Mass.: MIT Press, 150-156.

[35] Szymanski, S. (1991), The Optimal Timing of Infrastructure Investment, *Journal of Transport Economics and Policy*, **25**, 247-258.

[36] Tarola, O. (2006), Monopoly Pricing over Time and the Timing of Investments, *B.E. Journal of Theoretical Economics (Topics)*, **6**, 1-17.

[37] Tarola, O., and Trento, S. (2010), Timing of Investment and Dynamic Pricing in Privatized Sectors, *DISA Working Paper*, No. 1005, University of Trento, 1-15.

[38] Terborgh, G. (1949), *Dynamic Equipment Policy*, New York: McGraw-Hill.

[39] Tirole, J. (1988), *The Theory of Industrial Organization*, Cambridge, Mass: MIT Press.

[40] van Hilten, O. (1991), The Optimal Lifetime of Capital Equipment, *Journal of Economic Theory*, **55**, 449-451.

[41] Waterson, M. (1988), *Regulation of the Firm and Natural Monopoly*, Oxford: Basil Blackwell.

[42] Willig, R.D. (1976), Consumer's Surplus without Apology, *American Economic Review*, **66**, 589-597.

A Appendix

A.0.1 Proof of Proposition 1

Proof. The first-order conditions for a maximum of (10) with respect to Q_1 and Q_2 are:

$$\int_0^{T_I} \left((\gamma - 2)\, bQ_1^{DMM} + a - ce^{\alpha\left(t+\overline{T}\right)} \right) e^{-rt} dt = 0, \tag{25}$$

$$\int_{T_I}^{T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}} \left((\gamma - 2)\, bQ_2^{DMM} + a - ce^{\alpha(t-T_I)} \right) e^{-rt} dt = 0.$$

Thus, the time paths of outputs, Q_1^{DMM} and Q_2^{DMM}, which maximize a discounted convex combination of total surplus and profits for a given weighting factor, γ, before and after investment can be determined:

$$Q_1^{DMM} = \frac{a - ce^{\alpha\left(t+\overline{T}\right)}}{(2-\gamma)\, b},\ Q_2^{DMM} = \frac{a - ce^{\alpha(t-T_I)}}{(2-\gamma)\, b}. \tag{26}$$

The second-order conditions for a maximum of (10) with respect to Q_1 and Q_2 are fulfilled at Q_1^{DMM} and Q_2^{DMM}. Inserting Q_1^{DMM} and Q_2^{DMM} into (2) and taking into account that $g = 0$, it becomes obvious that these outputs result from the corresponding convex combinations of marginal cost prices and profit-maximizing prices at any point in time, i.e. $P\left(Q_1^{DMM}, t\right) = \frac{\gamma}{2-\gamma} ce^{\alpha\left(t+\overline{T}\right)} + \left(1 - \frac{\gamma}{2-\gamma}\right) \frac{a+ce^{\alpha\left(t+\overline{T}\right)}}{2}$ and $P\left(Q_2^{DMM}, t\right) = \frac{\gamma}{2-\gamma} ce^{\alpha(t-T_I)} + \left(1 - \frac{\gamma}{2-\gamma}\right) \frac{a+ce^{\alpha\left(t-T_I\right)}}{2}$. ■

A.0.2 Proof of Proposition 2

Proof. Inserting Q_1^{DMM} and Q_2^{DMM} from (26) into (10) yields:

$$\begin{aligned} DMM\left(T_I\right) = & \int_0^{T_I} \frac{\left(a-ce^{\alpha\left(t+\overline{T}\right)}\right)^2}{(4-2\gamma)b} e^{-rt} dt \\ & + \int_{T_I}^{T_I + \frac{\ln\left(\frac{a}{c}\right)}{\alpha}} \frac{\left(a-ce^{\alpha\left(t-T_I\right)}\right)^2}{(4-2\gamma)b} e^{-rt} dt \\ & -Ie^{-rT_I}. \end{aligned} \tag{27}$$

(27) can be written as:

$$\begin{aligned} DMM\left(T_I\right) = & -\tfrac{a^2}{(4-2\gamma)br}\left(e^{-rT_I} - 1\right) \\ & -\tfrac{ace^{\alpha\overline{T}}}{(2-\gamma)b(\alpha-r)}\left(e^{(\alpha-r)T_I} - 1\right) \\ & +\tfrac{c^2 e^{2\alpha\overline{T}}}{(4-2\gamma)b(2\alpha-r)}\left(e^{(2\alpha-r)T_I} - 1\right) \\ & -\tfrac{a^2}{(4-2\gamma)br}\left(e^{-r\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} - 1\right) e^{-rT_I} \\ & -\tfrac{ac}{(2-\gamma)b(\alpha-r)}\left(e^{(\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} - 1\right) e^{-rT_I} \\ & +\tfrac{c^2}{(4-2\gamma)b(2\alpha-r)}\left(e^{(2\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} - 1\right) e^{-rT_I} \\ & -Ie^{-rT_I}. \end{aligned} \tag{28}$$

The first-order condition for a maximum of (28) with respect to T_I is:

$$\begin{aligned}
\frac{dDMM}{dT_I} = \quad & \tfrac{a^2}{(4-2\gamma)b} e^{-rT_I^{DMM}} \\
& -\tfrac{ace^{\alpha\overline{T}}}{(2-\gamma)b} e^{(\alpha-r)T_I^{DMM}} \\
& +\tfrac{c^2 e^{2\alpha\overline{T}}}{(4-2\gamma)b} e^{(2\alpha-r)T_I^{DMM}} \\
& +\tfrac{a^2}{(4-2\gamma)b}\left(e^{-r\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right)e^{-rT_I^{DMM}} \\
& +\tfrac{acr}{(2-\gamma)b(\alpha-r)}\left(e^{(\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right)e^{-rT_I^{DMM}} \\
& -\tfrac{c^2 r}{(4-2\gamma)b(2\alpha-r)}\left(e^{(2\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}-1\right)e^{-rT_I^{DMM}} \\
& +rIe^{-rT_I^{DMM}} = 0.
\end{aligned} \tag{29}$$

This implies:

$$\begin{aligned}
\left(a-ce^{\alpha\overline{T}}e^{\alpha T_I^{DMM}}\right)^2 = \quad & a^2\left(1-e^{-r\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}\right) \\
& +\tfrac{2acr}{\alpha-r}\left(1-e^{(\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}\right) \\
& -\tfrac{c^2 r}{2\alpha-r}\left(1-e^{(2\alpha-r)\frac{\ln\left(\frac{a}{c}\right)}{\alpha}}\right) \\
& -(4-2\gamma)\,brI.
\end{aligned} \tag{30}$$

Using $e^{\alpha\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} = \frac{a}{c}$ and $e^{2\alpha\frac{\ln\left(\frac{a}{c}\right)}{\alpha}} = \frac{a^2}{c^2}$ we get:

$$\begin{aligned}
& \left(a-ce^{\alpha\overline{T}}e^{\alpha T_I^{DMM}}\right)^2 \\
= \quad & a^2+\tfrac{2acr}{\alpha-r}-\tfrac{c^2 r}{2\alpha-r}-\tfrac{\alpha^2 2a^2}{(r^2-3\alpha r+2\alpha^2)\left(\frac{a}{c}\right)^{\frac{r}{\alpha}}}-(4-2\gamma)\,brI.
\end{aligned} \tag{31}$$

Solving this quadratic equation for T_I^{DMM} leads to two potential candidates for an investment date maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ. However, only the one described by (11) is in the relevant range. It can be seen that the second-order condition for a maximum value of (28) at T_I^{DMM} is fulfilled.

$$\frac{d^2DMM}{d\left(T_I^{DMM}\right)^2} = -r\frac{dDMM}{dT_I^{DMM}} + e^{-rT_I^{DMM}}\frac{1}{b}c\alpha e^{\alpha\overline{T}}\frac{e^{\alpha T_I^{DMM}}}{\gamma-2}\left(a-ce^{\alpha\overline{T}}e^{\alpha T_I^{DMM}}\right), \tag{32}$$

and $\frac{d^2DMM}{d\left(T_I^{DMM}\right)^2}<0$, since $\frac{dDMM}{dT_I^{DMM}}=0$, and $a-ce^{\alpha\overline{T}}e^{\alpha T_I^{DMM}}>0$ because of (11). ■

A.0.3 Proof of Proposition 4

Proof. The first-order condition for a maximum of (20) with respect to Q_2 is:

$$\int_{T_I}^{T_2^{bind,DMM}} \left((\gamma - 2)\, be^{-gt} Q_2^{DMM} + a - c \right) e^{-rt} dt = 0. \tag{33}$$

Thus, the time path of outputs, Q_2^{DMM}, immediately after investment which maximize a discounted convex combination of total surplus and profits can be determined:

$$Q_2^{DMM} = \frac{(a-c)\, e^{gt}}{(2-\gamma)\, b}. \tag{34}$$

The second-order condition for a maximum of (20) with respect to Q_2 is fulfilled at Q_2^{DMM}. Inserting Q_2^{DMM} into (2) and taking into account that $\alpha = 0$, it becomes obvious that these outputs result from the corresponding convex combinations of marginal cost prices and profit-maximizing prices at any point in time, i.e. $P\left(Q_2^{DMM}, t\right) = \frac{\gamma}{2-\gamma} c + \left(1 - \frac{\gamma}{2-\gamma}\right) \frac{a+c}{2}$. ■

A.0.4 Proof of Proposition 5

Proof. Using (2) as well as the relationship $\alpha = 0$, and bearing in mind that a convex combination of marginal cost prices and profit-maximizing prices at any point in time is optimal, $T_1^{bind,DMM}$ and $T_2^{bind,DMM}$ can be found:

$$T_1^{bind,DMM} = \frac{\ln\left(\frac{(2-\gamma)bK_1}{a-c}\right)}{g},\ T_2^{bind,DMM} = \frac{\ln\left(\frac{(2-\gamma)bK_2}{a-c}\right)}{g}. \tag{35}$$

Inserting Q_2^{DMM} and the values of $T_1^{bind,DMM}$ and $T_2^{bind,DMM}$ into (20) yields:

$$\begin{aligned} DMM\left(T_I\right) = & \int_{\frac{\ln\left(\frac{(2-\gamma)bK_1}{a-c}\right)}{g}}^{T_I} \left(-\tfrac{1}{2}\left(2-\gamma\right) be^{-gt} K_1^2 + (a-c)\, K_1\right) e^{-rt} dt \\ & + \int_{T_I}^{\frac{\ln\left(\frac{(2-\gamma)bK_2}{a-c}\right)}{g}} \frac{(a-c)^2}{2(2-\gamma)b} e^{gt} e^{-rt} dt \\ & + \int_{\frac{\ln\left(\frac{(2-\gamma)bK_2}{a-c}\right)}{g}}^{\infty} \left(-\tfrac{1}{2}\left(2-\gamma\right) be^{-gt} K_2^2 + (a-c)\, K_2\right) e^{-rt} dt \\ & - Ie^{-rT_I}. \end{aligned} \tag{36}$$

The first-order condition for a maximum of (36) with respect to T_I is:

$$\begin{aligned} \frac{dDMM}{dT_I} = & -\tfrac{1}{2}(2-\gamma)\,bK_1^2 e^{-(g+r)T_I^{DMM}} + (a-c)\,K_1 e^{-rT_I^{DMM}} \\ & -\tfrac{(a-c)^2 e^{(g-r)T_I^{DMM}}}{2(2-\gamma)b} + rIe^{-rT_I^{DMM}} = 0. \end{aligned} \tag{37}$$

Solving this quadratic equation for T_I^{DMM} leads to two potential candidates for an investment date maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ. However, only the one described by (21) is in the relevant range. It can be seen that the second-order condition for a maximum value of (36) at T_I^{DMM} is fulfilled:

$$\begin{aligned} \frac{d^2DMM}{d\left(T_I^{DMM}\right)^2} = & -(g+r)\left(-\tfrac{1}{2}\right)(2-\gamma)\,bK_1^2 e^{-(g+r)T_I^{DMM}} \\ & -r\,(a-c)\,K_1 e^{-rT_I^{DMM}} - (g-r)\,\tfrac{(a-c)^2 e^{(g-r)T_I^{DMM}}}{2(2-\gamma)b} \\ & -r^2 Ie^{-rT_I^{DMM}} \\ = & -(g+r)\,\tfrac{dDMM}{dT_I^{DMM}} \\ & +g\,(a-c)\,K_1 e^{-rT_I^{DMM}} - g\tfrac{(a-c)^2 e^{(g-r)T_I^{DMM}}}{(2-\gamma)b} + grIe^{-rT_I^{DMM}}. \end{aligned} \tag{38}$$

Since $\frac{dDMM}{dT_I^{DMM}} = 0$, a sufficient condition for a maximum value of (36) at T_I^{DMM} is:

$$g\,(a-c)\,K_1 e^{-rT_I^{DMM}} - g\frac{(a-c)^2 e^{(g-r)T_I^{DMM}}}{(2-\gamma)\,b} + grIe^{-rT_I^{DMM}} < 0. \tag{39}$$

Rearranging (39) yields:

$$T_I^{DMM} > \frac{1}{g}\ln\left((2-\gamma)\,b\frac{K_1\,(a-c)+rI}{(a-c)^2}\right). \tag{40}$$

A comparison of (40) with (21) completes the proof. ■

A.0.5 Proof of Proposition 7

Proof. First, we focus on a monopoly firm maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ. Analogous to the derivation of Q_2^{DMM} in the Proof of Proposition 1, it can easily be seen that:

$$Q_{n+i}^{DMM} = \frac{a - ce^{\alpha(t-T_{n+i-1})}}{(2-\gamma)\,b}, \tag{41}$$

where $i \in \mathbb{N} \backslash \{0\}$. Inserting Q_{n+i}^{DMM} into (22) leads to:

$$\begin{aligned}
&DMM\left(T_n, T_{n+1}, T_{n+2}, T_{n+3}, \ldots\right) \\
&= \int_{T_n}^{T_{n+1}} \frac{\left(a - ce^{\alpha(t-T_n)}\right)^2}{(4-2\gamma)b} e^{-rt} dt - Ie^{-rT_{n+1}} \\
&+ \int_{T_{n+1}}^{T_{n+2}} \frac{\left(a - ce^{\alpha\left(t-T_{n+1}\right)}\right)^2}{(4-2\gamma)b} e^{-rt} dt - Ie^{-rT_{n+2}} \\
&+ \int_{T_{n+2}}^{T_{n+3}} \frac{\left(a - ce^{\alpha\left(t-T_{n+2}\right)}\right)^2}{(4-2\gamma)b} e^{-rt} dt - Ie^{-rT_{n+3}} \\
&\vdots
\end{aligned} \quad . \tag{42}$$

The first-order condition for a maximum value of (42) with respect to T_{n+1} can be described as:

$$\begin{aligned}
\frac{\partial DMM}{\partial T_{n+1}} = \; & \frac{\left(a - ce^{\alpha\left(T_{n+1}^{DMM} - T_n^{DMM}\right)}\right)^2 e^{-rT_{n+1}^{DMM}}}{(4-2\gamma)b} + rIe^{-rT_{n+1}^{DMM}} - \frac{(a-c)^2}{(4-2\gamma)b} e^{-rT_{n+1}^{DMM}} \\
& + \frac{\alpha}{\alpha - r} \frac{ac}{(2-\gamma)b} e^{-\alpha T_{n+1}^{DMM}} \left(e^{(\alpha-r)T_{n+2}^{DMM}} - e^{(\alpha-r)T_{n+1}^{DMM}}\right) \\
& - \frac{\alpha}{2\alpha - r} \frac{c^2 e^{-2\alpha T_{n+1}^{DMM}}}{(2-\gamma)b} \left(e^{(2\alpha-r)T_{n+2}^{DMM}} - e^{(2\alpha-r)T_{n+1}^{DMM}}\right) = 0.
\end{aligned} \tag{43}$$

(43) is equivalent to:

$$\begin{aligned}
& \left(a - ce^{\alpha\left(T_{n+1}^{DMM} - T_n^{DMM}\right)}\right)^2 + 2\left(2-\gamma\right) brI - \left(a-c\right)^2 \\
& + \frac{\alpha}{\alpha - r} 2ace^{(\alpha-r)\left(T_{n+2}^{DMM} - T_{n+1}^{DMM}\right)} - \frac{\alpha}{\alpha - r} 2ac \\
& - \frac{\alpha}{2\alpha - r} 2c^2 e^{(2\alpha-r)\left(T_{n+2}^{DMM} - T_{n+1}^{DMM}\right)} + \frac{\alpha}{2\alpha - r} 2c^2 = 0.
\end{aligned} \tag{44}$$

This is an infinite replacement problem, where replacement conditions do not change over time. Therefore, the interval, x_{DMM}, between any two successive investment dates, T_{n+i}^{DMM} and T_{n+i-1}^{DMM}, maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ, is constant:

$$x_{DMM} = T_{n+i}^{DMM} - T_{n+i-1}^{DMM}, \tag{45}$$

where $i \in \mathbb{N} \backslash \{0\}$.

Using (45) we obtain:

$$\begin{aligned}
rI = \; & \frac{1}{(4-2\gamma)b} \Big[- \left(a - ce^{\alpha x_{DMM}}\right)^2 + \left(a-c\right)^2 - \frac{\alpha}{\alpha - r} 2ace^{(\alpha-r)x_{DMM}} \\
& + \frac{\alpha}{\alpha - r} 2ac + \frac{\alpha}{2\alpha - r} 2c^2 e^{(2\alpha-r)x_{DMM}} - \frac{\alpha}{2\alpha - r} 2c^2 \Big].
\end{aligned} \tag{46}$$

Denoting the term in squared brackets $f(\cdot)$, which is a function of x_{DMM}, yields:

$$rI = \frac{1}{(4-2\gamma)\,b} f(x_{DMM}) \,. \tag{47}$$

Since $f(0) = 0$, it is obvious that $\frac{1}{(4-2\gamma)b} f(0) < rI$. Furthermore, $\frac{df(x_{DMM})}{dx_{DMM}} > 0$ for $x_{DMM} < \frac{\ln\left(\frac{a}{c}\right)}{\alpha}$. Thus, we can infer that there is exactly one solution for (47) in the relevant range, if an equilibrium exists.

Now, we need to distinguish between three cases. Either $\gamma = 0$, or $0 < \gamma < 1$, or $\gamma = 1$. Insertion into (47) and comparison of the three cases conclude the proof. ■

A.0.6 Proof of Proposition 8

Proof. Initially, we focus on a monopoly firm maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ. Analogous to the derivation of Q_2^{DMM} in the Proof of Proposition 4, it can easily be seen that:

$$Q_{n+i}^{DMM} = \frac{(a-c)\,e^{gt}}{(2-\gamma)\,b}, \tag{48}$$

where $i \in \mathbb{N}$.

Furthermore, it is obvious that:

$$T_{n+i}^{bind,DMM} = \frac{\ln\left(\frac{(2-\gamma)bK_{n+i+1}}{a-c}\right)}{g}, \tag{49}$$

where $i \in \mathbb{N}$.

Inserting Q_{n+i}^{DMM} and the values of $T_{n+i}^{bind,DMM}$ into (24) yields:

$$
\begin{aligned}
&DMM\left(T_n, T_{n+1}, T_{n+2}, T_{n+3}, \ldots\right) \\
&= \tfrac{(a-c)^2}{(4-2\gamma)b} \int\limits_{T_n}^{\frac{\ln\left(\frac{(2-\gamma)bK_{n+1}}{a-c}\right)}{g}} e^{(g-r)t}dt \\
&+ \int\limits_{\frac{\ln\left(\frac{(2-\gamma)bK_{n+1}}{a-c}\right)}{g}}^{T_{n+1}} \left(\left(\tfrac{\gamma}{2}-1\right)be^{-gt}K_{n+1}^2 + (a-c)K_{n+1}\right)e^{-rt}dt - Ie^{-rT_{n+1}} \\
&+ \tfrac{(a-c)^2}{(4-2\gamma)b} \int\limits_{T_{n+1}}^{\frac{\ln\left(\frac{(2-\gamma)bK_{n+2}}{a-c}\right)}{g}} e^{(g-r)t}dt \\
&+ \int\limits_{\frac{\ln\left(\frac{(2-\gamma)bK_{n+2}}{a-c}\right)}{g}}^{T_{n+2}} \left(\left(\tfrac{\gamma}{2}-1\right)be^{-gt}K_{n+2}^2 + (a-c)K_{n+2}\right)e^{-rt}dt - Ie^{-rT_{n+2}} \\
&\vdots
\end{aligned}
\tag{50}
$$

The first-order condition for a maximum of (50) with respect to T_{n+1} can be described as:

$$
\begin{aligned}
\tfrac{\partial DMM}{\partial T_{n+1}} = \quad & \left(\tfrac{\gamma}{2}-1\right)bK_{n+1}^2 e^{-(g+r)T_{n+1}^{DMM}} + (a-c)K_{n+1}e^{-rT_{n+1}^{DMM}} \\
& + rIe^{-rT_{n+1}^{DMM}} - \tfrac{(a-c)^2}{(4-2\gamma)b}e^{(g-r)T_{n+1}^{DMM}} = 0.
\end{aligned}
\tag{51}
$$

Solving this equation for T_{n+1}^{DMM} leads to two potential candidates for an investment date maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ. However, only the following one is in the relevant range:

$$
\begin{aligned}
T_{n+1}^{DMM} = \quad & \tfrac{1}{g}\ln\left[\tfrac{(2-\gamma)b}{(a-c)^2}\left((a-c)K_{n+1} + rI\right.\right. \\
& \left.\left. + \sqrt{\left((a-c)K_{n+1} + rI\right)^2 - K_{n+1}^2(a-c)^2}\right)\right].
\end{aligned}
\tag{52}
$$

Now, it becomes obvious that the time interval between two successive investments for a monopoly firm maximizing a discounted convex combination of social welfare and profits for a given weighting factor, γ, is not constant over time, i.e. $x_{n+1,n}^{DMM} \neq x_{n+2,n+1}^{DMM} \neq x_{n+3,n+2}^{DMM}$... Therefore, a meaningful comparison of identical monopoly firms with different objectives is only possible for time intervals between *corresponding* successive investments, i.e. only for *those with identical subscripts*. We can only compare $x_{n+1,n}^{DTS}$ to $x_{n+1,n}^{DMM}$ and to $x_{n+1,n}^{D\Pi}$.

To do this, we need to distinguish between three cases. Either $\gamma = 0$, or $0 < \gamma < 1$, or $\gamma = 1$. Insertion into (52) and comparison of the time intervals between corresponding successive investments for the three cases conclude the proof. ■