Essays on R&D, investment and growth

Ornella Tarola

Université catholique de Louvain

Faculté des sciences économiques, sociales et politiques

Département des sciences économiques

Supervisor: Prof. R. Boucekkine

Co-Supervisor: Prof. G.Cozzi

Contents

Acknowledgements

I thank professor Raouf Boucekkine, without whom writing this thesis would have been impossible. Also, I thank professor Guido Cozzi –coauthor of chapter 1- for helping me to understand and appreciate Growth Theory, and Stefano Demichelis –coauthor of chapter 2- for his fundamental contribution to chapter 2. I am grateful to professor David de la Croix, whose comments contributed to ameliorate considerably this thesis. I am indebted to professor Jean J.Gabszewicz for his sharp suggestions, and crucial discussions on this thesis and different issues of Economic Theory. I also would like to thank Emanuele Bacchiega - my coauthor -for his conscientiousness while working, professor Giovanni Caravale and professor Giorgio Rodano for fueling my interests in economics, the Department of Economics, and people working at CORE and IRES for their valuable help. I am also grateful to the Institute of Economic Research, University of Lugano, whose support allowed me to finish the redaction of this work.

A thank to my friends: they were able to support me in spite of the geographical distance between us.

A special thank to my parents and my husband, for making my life worth living every day.

Introduction

In the late 1980's and throughout 1990's the investment policy in the R&D sector became progressively cooperative, and research joint ventures proliferated. Indeed, after positive results of large R&D agreements in US and Japan, even European firms started to adopt cooperative investment policies in R&D. As a consequence of this raising phenomenon of cooperative policies, a large body of industrial organization literature (d'Aspremont and Jacquemin 1988, Choi 1989, Kamien et al 1992, *inter alia*) examined motives for firms to pool their R&D investments. It was found that knowledge generated by research labs is often valuable to strongly different firms, and alliances in R&D reduce the unnecessary duplication of efforts which occurs when independent research labs invest in identical or alike research projects, and enable firms to take a *synergistic advantage* through sharing information. Yet, the impact of R&D cooperation on innovative activity turned out to be quite ambiguous, this being dependent on the spillovers' magnitude among firms. In d'Aspremont and Jacquemin (1988) and Choi (1989), cooperative R&D determines superior technological advancement with respect to competitive R&D only if the spillovers are big enough, while in Kamien Muller and Zang (1992) conducting R&D cooperatively always determines better advancement that conducting R&D competitively. Moreover, while considering the relation between cooperative investment policy in R&D and innovative activity, the industrial organization literature disregarded to analyise the impact of a cooperative investment policy in R&D on the growth rate of the economy. Surprisingly, scarce attention has been also paid by endogenous growth theory to this issue (Cozzi 1999, Link, Paton and Siegel 2002, Shapiro 2002): neither incentives for firms to conduct cooperative investment policy in the R&D sector, nor the main impact this policy could have on the growth path of the economy have been evaluated by the growth literature, although it has been recognized that strategic agreements between research units may strongly affect innovative activity and thus growth(Jovanovic and Rob 1989, Jovanovic and MacDonald 1994, Jovanovic and Rousseau 2001).

Thus, taking into account the above mentioned idea of synergistic advantage as it has been developed in the partial equilibrium approach of industrial organization, and the surprising gap in the growth theory, we embed this issue into an endogenous R&D-based growth model. We start considering the economy as a whole in order to evaluate how its actual functioning (interaction between research labs, risk of failure in innovating, patent laws and so on) shapes firms' specific investment policies. More precisely, we focus on the R&D process as it really develops in the market, and examine what is the optimal investment policy for firms involved in research and whether the selected investment policy in turn sustains the growth path of the economy. The main questions this model should answer are whether it is profitable for firms involved in a R&D race to cooperate in their own investment policy, through sharing information, mergers and so on, and how this investment policy, as it is actually conducted, affects the innovative activity and thus the growth path of the economy as a whole.

Then, after analysing how the economic environment may affect firms' investment plans, we propose to identify the decision mechanism, as it develops inside the firm, through which an investment policy is defined over time. Accordingly, we move to an in-depth analysis of the process through which a profit-maximizing entity selects a specific investment policy when it is required to expand its production plant in order to satisfy a growing economy. For our purpose, we turn to an old trend of literature related to the so called plant size problem and developed since 1950's by Chanery (1952) and Manne (1961). Most of it has been conceived and formulated for solving planning problems met by governments in developing countries (surveys on the main applications of capacity expansion models are provided by Luss 1982, Li and Tirupati 1992, and Nam and Logendran 1992), where demand expansion was pushed exogenously by the population growth (Bean and al. 1992, Chaouch and Buzacott 1994, and Ryan 2002, *inter alia*). Although it is no longer reasonable to treat firms as public institutions, and the globalization of trade has replaced demographic booms in pushing up demand, the plant size literature can still represent an useful starting point for analyzing the rationale on which investment policies lie. The basic issue of the plant size approach comes from the co-existence, in some sectors of the economy, of important economies-of-scale combined with a significant expansion of demand. Delaying investment is then beneficial for two reasons. First, the present value of the resulting costs is lower and, second, the increase in demand becoming larger and larger over time, allows to benefit better from the economies of scale. Accounting for this and accordingly within the plant size modelling, we investigate the criterion for investing in a growing economy when firms are no longer public entities, as in the original situation, but rather profit maximizing institutions. As such, we assume that they can accompany their investment policy by a price policy through which to manipulate the level of demand through time. Then, both of these policies can be conducted in order to put in balance costs and benefits which derive from investing. The main finding provided by this literature since Manne (1961) was a constant-cycle property in the firm' s optimal investment policy, according to which the firm installs a new facility of equal size at equally-spaced dates. We examine the characteristics of the optimal policy of the monopolist in terms of the periods of time during which the production capacity can be expanded and verify if this constant-cycle property still holds when the demand function can be shaped through time.

The first issue of how to conduct investment policies in research taking into account the actual functioning of the economy as a whole is extensively examined in the first part of this work, while the decision mechanism through which the investment policy is planned by a firm is considered in the remaining part of the work. In the following, we provide a brief description of the content of each analysis and discuss further steps of research which could derive from them.

Mergers, Innovation, and Inequality

In the late 1980's and throughout 1990's a relevant increase in R&D cooperation was observed, and research joint ventures (RJV) proliferated. While recognizing that strategic agreements between research units may strongly affect innovative activity and thus growth, scarce attention has been paid by endogenous growth theory to this issue (Cozzi, 1999). On the other hand there is a large body of industrial organization literature (d'Aspremont and Jacquemin 1988, Choi 1989, Kamien et al. 1992, *inter alia*) which examines motives for firms to pool their R&D efforts. It is commonly argued that knowledge generated by research labs is often valuable to strongly different firms, while alliances in R&D reduce the unnecessary duplication of efforts which occurs when independent research labs invest in identical or alike research projects, and enable firms to take a synergistic advantage through sharing information. Typically, in a context of a two stage game, firms are assumed to conduct research in the first stage and engage in either Bertrand or Cournot competition in the product market in the second stage. When firms conduct cooperative R&D, their efforts result in cheaper or better advancement owing to the internalization of technological spillovers, namely sharing knowledge and useful information, whereas the impact of R&D cooperation on innovative activity is quite ambiguous, this being dependent on the spillovers' magnitude among firms. In d'Aspremont and Jacquemin (1988) and Choi (1989), cooperative R&D determines superior technological advancement with respect to competitive R&D only if the spillovers are big enough, while in Kamien Muller and Zang (1992) conducting R&D cooperatively always determines better advancement than conducting R&D competitively. Although this literature examines exactly the relation between R&D cooperation and innovative activity, it is traditionally concerned with a partial-equilibrium approach.

In this paper we incorporate the issue developed by the industrial organization literature to an endogenous growth model. We address both the questions examined by industrial organization models, analysing what induces firms to cooperate in research and how R&D cooperation affects innovative activity and growth. We assume that research activity is an inventing lottery: an infinite number of monopolistically competitive firms are involved in a costly and stochastic R&D race; only who first markets an innovation is rewarded with a patent, the rivals being temporarily out of the market. While this is a traditional premise of recent Schumpeterian growth models, the innovation process is deterministic in the industrial organization approach and the patent protection is usually considered in a separate frame. Accounting for this inventing lottery enables us to examine very realistically how R&D cooperation develops when firms are uncertain whether they will win the R&D race and will be rewarded for their purposive research effort. Moreover, we show how cooperative R&D affects research activity and growth regardless of the benefits' magnitude that firms may exploit by pooling their research efforts.

Our model displays also several properties which are interesting in growth theory. We avoid the well-known scale effect suffered by the most part of endogenous growth models and which has been proved by Jones (1995) to be at odd with the empirical evidence. Recently, there have been many attempts to overcome this prediction suffered by earlier models (Grossman and Helpman 1991, Aghion and Howitt 1992 and 1998, Jones 1995, Kortum 1997, Dinopoulos and Segerstrom 1998, Young 1998, Peretto 1998, Segerstom 1998, Howitt 1999). Although they represent important advancements in removing the scale effect property, Jones (1999) shows that their results are sensitive to the value of parameters at the basis of the analysis. Here, we provide a new rationale for avoiding the scale effect, without incurring in the criticism formulated by Jones (1999). Further, our analysis embeds an issue long considered in growth theory, although confined to a separate set of models (Greenwood, Hercowitz and Krussel 1997, Greenwood and Yorucoglu 1997, Allen 2001), namely the relationship between technological advancement and wage inequality occurred since the late 1970's in the US. We illustrate how these phenomena are linked and provide a new explanation to the radical shift in the distribution of wages and the high skill premium which emerged when the technological skill-biased revolution took place (Kremer and Maskin 1999, Acemoglu 1999, 2002).

Our results seem consistent with some well known empirical evidence on inter-firm barriers, the related incentives for firms to cooperate, and the plausible link between market structure, wage inequality, and growth(Agarwal and Gort 2000). Neverthless, they have been obtained within the extremely stylized Aghion and Howitt's (1992) framework. It would be interesting to extend our logic to an explicitly more complex framework, such as for example Aghion et al. (2001). Furthermore, by analyzing US data in the period 1985-1998 Link, Paton and Siegel (2002) find that authorities systematically allow R&D agreements more often when the country is losing international competitiveness in high technology sectors. While this finding is consistent with our claim that policy changes are at the root of the observed consequence for innovation and income distribution, at the same time it suggests to endogenize industrial policy in an explicit international setting. These will be topics for further work.

Capacity Expansion and Dynamic Monopoly Pricing

The so-called plant-size problem introduced in the fifties by Chenery (1952) studies the optimal capacity to install in order to serve an exogenous demand increasing over time, with economies of scale in plant construction. This problem arose in the context of developing countries: explosive demographics forced the governments to regularly increase capacity over time, in order to face increase in demand resulting from such populations' booms. Increasing returns to scale in plant construction opened the door to the possibility of anticipating the increments of capacity required to meet future demand levels, and exploiting thereby the resulting economies of scale. Manne (1961) and (1967) starting from the capacity expansion problem which was met by several manufacturing facilities in India, examined the rationale behind the optimal investment policy faced by public institutions. Assuming an exogenous demand growing linearly over time and economies of scale in investment costs, he finds that the optimal investment policy is constant-cycle : successive investments are all of the same size and undertaken at equally spaced points of time.

While in this context, it was reasonable to assume that demand expansion was pushed exogenously by the population growth, today, in developed countries, similar demand booms are still observed in several industries, but for other reasons than an exogenous population growth. Indeed, markets in developed countries expanded initially around local or national demands, whereas now they are progressively concerned by larger and larger geographical areas. In these areas, firms do not act as public institutions, but rather as profit maximizing companies. As such, they accompany their investment policy by a price policy through which they can manipulate the levels of demand through time. Although the seminal paper by Manne has been extended in several directions (Srinivasan 1967, Gabszewicz and Vial 1972, Nickell 1977, Freidenfelds 1981, Bean and al. 1992, Chaouch and Buzacott 1994, and Ryan 2002 *inter alia*), these research contributions assume that demand for capacity increases is exogenously given, and cannot be adjusted through some price policy.

In order to catch the strategic possibility for a profit-maximizing firm to manipulate the demand function through price, we study in the present paper a similar plant size problem as in Chenery (1952) and Manne (1961). Yet we now allow the profit-maximising firm producing the good to combine its investment policy with a product price policy adjusting demand upwards, or downwards, over time. More precisely, we suppose that, at each point of time, market demand is given by a linear function of instantaneous price, the intercept of which increases linearly with time. Moreover, we make the assumption alluded to above concerning the size of the investments, namely, we fix a sequence of equally spaced points of time at which the investments for adding capacity may be undertaken and characterize the optimal policy of the monopolist in terms of price regime and investment's size. This extension has to cope with the difficulty that, even if the volume demanded is exogenously increasing (or decreasing) at each price over time, the firm can now decrease or increase it by instantaneous increases or decreases in the price pattern. We characterize the optimal price and investment policies of the monopolist under the assumption that the firm while controlling the size of the investments to be undertaken, does not control the dates, which are assumed to be equally spaced through time, at which these investments have to be consented. Our findings are as follows. The optimal price policy, through which the monopolist can dampen or enhance the expansion of demand through time, leads to an investment pattern with increments of capacity which are either constant over time at each investment date, or start to become constant after a finite set of investment dates. Furthermore, the optimal price pattern between two dates at which new capacity is installed may follow two different directions: either the price pattern is such that total existing capacity is fully used at each instant between these dates; or the instantaneous monopoly price is used for some period within the cycle and, thereafter, until the end of the cycle, the price which dampens instantaneous demand at the level of installed capacity is adopted. Finally, we show that the optimal constant increment of capacity is smaller than the one which would have been selected by a planner facing a demand growing exogenously at a rate corresponding to the instantaneous monopoly price, as in Manne (1961).

Monopoly Pricing over Time and the Timing of Investments

Adding facilities to optimally meet rising demand for these facilities is one of the most debated issue in all the industries which are characterized by trade-offs in decision making process due to economies of scale in investment costs or commitment of substantial resources for further capacity. Typical examples are given by heavy process industries, communication network, and water resources systems. Research on this topic has been pursued since the sixties in the context of developing countries where explosive demographics forced the governments to regularly increase capacity over time, in order to face increases in demand resulting from such populations' booms. Assuming an exogenous demand growing linearly over time and economies of scale in investment costs, Manne (1961 and 1967) in his pioneristic work, found that the optimal investment policy is constant-cycle: successive investments are all of the same size and undertaken at equally spaced points of time. These findings were extended in several directions (Srinivasan 1967, D'Aspremont Gabszewicz and Vial 1972, Gabszewicz and Vial 1972), and variants of the capacity expansion model as it was formulated by Manne (1961) were used for several application to public services (see Luss 1982, Li and Tirupati 1992, and Nam and Logendran 1992 for a survey of the capacity expansion problem in OR and its main applications).

While in the context of developing countries, it was reasonable to assume that demand expansion was pushed exogenously by the population growth, today similar demand booms are observed for other reasons than an exogenous population growth. Indeed, they mainly derive from the globalisation of trade which makes industries progressively concerned with larger and larger geographical areas to be served.

Further, mainly in response to the failure of state-owned entities to meet the growing needs of customers, many public utilities have been progressively changed into private institutions, namely profit maximizing entities. As such, they are allowed to use according to a profit-maximization criterion specific instruments which were not available at the time when public utilities were required to satisfy the demand function, e.g. they can adopt a price policy in order to manipulate the demand function over time consistently with their installed capacity.

Accordingly, we propose to relax the assumption of exogenous demand increase and replace it by an assumption about the time expansion of a demand function specifying how instantaneous demand varies with price at each point of time. More precisely, we consider a monopoly producing a good whose intercept of its demand function grows linearly over time: at each price demand expands through time as the intercept of the demand function grows. Whatever the price policy that has been adopted at each instant, the monopolist must invest in order to meet the resulting demand growth. Further, the monopolist can only invest in indivisible factory units. The problem we would like to study in this paper is the following. *What is the optimal price policy, and the ensuing optimal sequence of investment time points the monopolist should select through time ?* This is nothing else than the so-called capacity expansion problem formulated by Chenery (1952) and Manne (1961). Yet, now it examined the timing of capacity additions which is defined on the basis of a price manipulation, when the expansion capacity consists of a new factory.

Our findings are as follows. The optimal price policy, through which the monopolist can possibly dampen or enhance the expansion of demand through time, leads to a sequence of time points when to install a new factory which are equally spaced. Furthermore, the optimal price pattern between two dates at which a new factory is installed falls into two categories: either the price pattern is such that total existing capacity which results from the number of installed factories is fully used at each instant between these dates; or the instantaneous monopoly price is used for some period within these dates and, thereafter, until the point of time when a new factory is installed, the price which dampens instantaneous demand at the level of installed factory is adopted. Finally, we show that the optimal sequence of time points when new factories are installed determines cycles which are longer than the ones which would have been selected by a planner facing a demand growing exogenously at a rate corresponding to the instantaneous monopoly price, as in Manne (1961). This paper is also related to other bulks of literature. First, it is concerned with the application of microeconomic models of discontinuous adjustments (retail inventories, expansion of capacity, cash balances, prices, durable goods etc.) to the aggregate economic behavior (Abel and Eberly 1996, Caplin and Spulber 1987, Caballero and Engel 1991 and 1999, *inter alia*). In particular, on the basis of Caplin and Spulber (1987), Caballero and Engel (1991) have developed a methodology which allows to explain why the economic behaviour of the aggregate investment displays a smooth pattern at the aggregate level, while individual production units adjusting discretely and by large amount. Although this approach has been conceived within the framework of (S,s) policy, it can still be applied, with minor changes, to our model and thus used for evaluating macroeconomic implications of investment policy in terms of capacity expansion as it is defined at a plant level. Second, the stationarity property we find also holds in models concerned with the optimal timing of replacement investments. These replacement models assume that the technological progress is embodied in new equipment and accordingly examine when it is optimal replacing the oldest capital with the most recent vintages. The endogenous scrapping time, namely the age of the oldest machines, which is determined by the replacement timing selected by firms, is proved to be constant over time, under several different conditions (Malcomson 1975, Van Hilten 1991, Boucekkine, Germain and Licandro 1997, Boucekkine del Rio and Licandro 1999). Although our specification, as it is, does not focus on the replacement problem, our modeling fits into the literature on replacement decision. With slightly changes, we could assume that a factory is required to replace its equipment which deteriorates with new machines rather than to expand its capacity to meet a rising demand, and then evaluating at the optimal price policy the ensuing timing of replacement.

The rationale we develop in this model can easily be applied to the inventory problem when the firm is allowed to manipulate the selling price in order to accelerate or dampen the demand through its price policy at the exit of the warehouse.

Finally, it could be interesting to apply this model to industries where a privatization process has taken place and planning investments becomes crucial for firms to survive (e.g. air travel industry).

CHAPTER 1.

Mergers, Innovation, and Inequality

Guido Cozzi and Ornella Tarola

Abstract.

This paper presents a standard endogenous growth framework in which the source of growth is represented by vertical innovation. The crucial assumption we introduce is that there is a positive information gap concerning the discovery of innovation. The aim of reducing the information dissemination lag provides incentives for firms to decide to merge their research efforts. At the same time we find that the skilled/unskilled wage gap is strongly related to this phenomenon. We prove that changing antitrust attitudes toward efficienc-motivated mergers in contestable industries may simultaneously explain observed changes in the industry structure, in qualitative innovation, in wage inequality, and in labor supply composition.

1.1 Introduction

In the late 1980's and throughout 1990's a relevant increase in R&D cooperation was observed, and research joint ventures (RJV) proliferated. The positive results of large R&D agreements in US and Japan (such as the VLSI Circuit Association in Japan and SEMATECH in the US) stimulated even European firm to conduct cooperative R&D. While recognizing that strategic agreement between research units may strongly affect innovative activity and thus growth, scarce attention has been paid by endogenous growth theory to this issue (Cozzi, 1999). On the other hand there is a large body of industrial organization literature (d'Aspremont and Jacquemin 1988, Choi 1989, Kamien et al. 1992, inter alia) which examines motives for firms to pool their R&D efforts. It is commonly argued that knowledge generated by research labs is often valuable to strongly different firms, while alliances in R&D reduce the unnecessary duplication of effort which occurs when independent research labs invest in identical or alike research projects, and enable firms to take a synergistic advantage through sharing information. "The alleged advantage of a research joint ventures, aside from enabling the participants to overcome a cost-of-development barrier impenetrable to any one of them alone, is the elimination of duplication effort..." (Kamien Muller and Zang, 1992, p.1293). Tipically, in a context of a two stage game, firms are assumed to conduct research in the first stage and engage in either Bertrand or Cournot competition in the product market in the second stage. When firms conduct cooperative R&D, their efforts result in cheaper or better advancement owing to the internalization of technological spillovers, namely sharing knowledge and useful information, whereas the impact of R&D cooperation on innovative activity is quite ambiguous, this being dependent on the spillovers'magnitude among firms. In d'Aspremont and Jacquemin (1988) and Choi (1989), cooperative R&D determines superior technological advancement with respect to competitive R&D only if the spillovers are big enough, while in Kamien Muller and Zang (1992) conducting R&D cooperatively always determines better advancement than conducting R&D competitively. Although this literature examines exactly the relation between R&D cooperation and innovative activity, it is traditionally concerned with a partial-equilibrium approach.

In this paper we incorporate the issue developed by the industrial organization literature to an endogenous growth model. We address both the questions examined by industrial organization models, analysing what induces firms to cooperate in research and how R&D cooperation affects innovative activity and growth. We assume that research activity is an inventing lottery: an infinite number of monopolistically competitive firms are involved in a costly and stochastic R&D race; only who first markets an innovation is rewarded with a patent, the rivals being temporarily out of the market. While this is a traditional premise of recent Schumpeterian growth models, the innovation process is deterministic in the industrial organization approach and the patent protection is usually considered in a separate frame. Accounting for this inventing lottery enables us to examine very realistically how R&D cooperation develops when firms are uncertain whether they will win the R&D race and will be rewarded for their purposive research effort. Moreover, we show how cooperative R&D affects research activity and growth regardless of the benefits'magnitude that firms may exploit by pooling their research efforts.

Our model displays also several properties which are interesting in growth theory. We avoid the well-known scale effect suffered by the most part of endogenous growth models and which has been proved by Jones (1995) to be at odd with the empirical evidence. Recently, there have been many attempts to overcome this prediction suffered by earlier models (Grossman and Helpman 1991, Aghion and Howitt 1992, etc.). Jones (1995), Kortum (1997) and Segerstom (1998) assume that the arrival rate of innovation decreases with level of knowledge; while Howitt (1999), Dinopoulos and Segerstrom (1998), Young (1998), and Peretto (1998) consider horizontal and vertical innovations jointly. Although they represent important advancements in removing the scale effect property, Jones (1999) shows that their results are sensitive to the value of parameters at the basis of the analysis. Here, we provide a new rationale for avoiding the scale effect, without incurring in the criticism formulated by Jones (1999). Further, our analysis embeds an issue long considered in growth theory, although confined to a separate set of models (Greenwood, Hercowitz and Krussel 1997, Greenwood and Yorucoglu 1997, Allen 2001), namely the relationship between technological advancement and wage inequality occurred since the late 1970's in the US. We illustrate how these phenomena are linked and provide a new explanation to the radical shift in the distribution of wages and the high skill premium which emerged when the technological skill-biased revolution took place (Kremer and Maskin 1999, Acemoglu 1999, 2002).

We analyse the innovation activity in an endogenous growth model sharing several features with the well-known Aghion and Howitt's (1992 and 1998) basic

Schumpeterian model, in which the source of growth is represented by vertical (product quality improving) innovation. We distinguish between the research process addressed to invent new products (*discovery*) and the process which makes these inventions marketable (*commercialization*). While the discovery is determined by a stochastic process, the commercialization of new products occurs when a patent has been granted to the inventor of this novel, in order to prevent rivals from using, producing and selling the invention. When a patent application is filed, a patent examiner, in determining patentability, must find that the invention is useful, novel, and non-obvious and this examination process takes time. As no disclosure is required before getting a patent, the time when the discovery of innovation is made by a first firm is unknown by the other firms up to the time when a patent is granted. Thus, there is a period of time when rival firms do unnecessary research effort. This information-gap determines a duplication of effort's risk (usually mentioned in the industrial organization literature) and makes R&D cooperation profitable. Further, as the probability of a firm's winning the R&D race decreases with the number of firms in the market, cooperative R&D always makes more valuable innovating and increases, *ceteris paribus*, the wage gap between skilled and unskilled labor. Moreover, it leads to better technological advancements than competitive R&D, regardless of the duration of patent examination which only affects the magnitude of the benefits for firms to pool efforts. Finally, this information dissemination lag puts an upper bound to the growth rate of income removing the scale effect property displayed by Aghion and Howitt (1992).

1.2 The model

The economy is populated by a continuum of infinitely-lived individuals with linear intertemporal preferences

$$u(y) = \int_0^\infty y_\tau e^{-r\tau} d\tau$$

where r is a constant rate of time preference.

Due to perfect capital markets and linear preferences, the equilibrium interest rate will be constant and equal to r.

Each individual is endowed with a unit flow of labor, either unskilled labor M or skilled labor L.

Unskilled labor is only used to produce the final consumption good according to the following production function

$$y = F(M, x) = AM^{1-\alpha} x^\alpha$$

where A is a productivity parameter and x the intermediate inputs. As M is used in a fixed amount, it can be normalized to one.

Skilled labor is devoted to the innovation process. This process consists in *discovering* and *producing* new inputs whose use increases the technology parameter by γ.

The discovery process is modeled as a R&D race. A flow of skilled labor, say n, competes in research-labs for inventing new types of intermediate goods. Inventions $t = 0, 1...$ occur with a Poisson arrival rate λn where λ represents the productivity of research technology. The one who first invents a new intermediate good obtains after a fixed time δ a patent on the newest input. Let τ_t be the discovery time of inputs t, then $\tau_t + \delta$ is the time when these new inputs become marketable.

Between τ_t and $\tau_t + \delta$ no profit is earned and the invention is not disclosed by the successful innovator. Thus, during this time competitors of the unknown innovator still seek to discover an *existing,* although not *disclosed,* generation t of inputs.

Assume that the R&D race for inventing the generation t of inputs starts at time 0. At some time $\tau > 0$ no invention occurred yet.

If $\tau \geq \delta - \tau_t$, then the time when new inputs could have been invented is just δ, otherwise the invention would have been already marketed.

If $\delta > \tau - \tau_t$, namely the time for obtaining a patent has not elapsed since the R&D race started, then the time when the new type of inputs has been invented while not patented, is $(\tau - \tau_t)$.

Let τ_δ be the time when new type of inputs could have been discovered and not yet disclosed. Then, the following holds:

$$\tau_\delta = \min\{\delta, \tau - \tau_t\}$$

where τ is any istant of time after the discovery at time τ_t of inputs t.

As the innovating follows a Poisson process, the distribution of this random variable τ_δ is:

$$F(\tau_\delta) = 1 - e^{-\lambda n \tau_\delta}$$

Thereinafter we focus on $\tau_\delta = \delta$.

Once discovered and patented, new inputs are produced by skilled labor in an intermediate sector according to a one-for-one relation and sold to the final output sector. Accounting for the two competing uses of skilled labor, the skilled-labor market equilibrium equation writes:

$$L_t = x_t + n_t \tag{1}$$

The innovator having an exclusive right on the newest intermediate input, due to the patent can monopolize the intermediate sector and quote a monopoly price while selling intermediate goods to the final output sector.

Then, the profit accruing to this monopolist is as follows:

$$\pi_t = \max(p_t - w_t)x_t$$

where the wage rate of skilled labor w_t used for producing inputs x_t is taken as given. The wage rate of unskilled workers is given by w^u.

As the final output is assumed to be perfectly competitive, the price p_t of the intermediate input x_t must be equal to its marginal productivity, namely

$$p_t = A_t \alpha x^{\alpha-1}$$

Then, the intermediate monopolist's profit maximization problem yields the usual expression:

$$\pi_t = \left(\frac{1-\alpha}{\alpha}\right) w_t x_t \equiv A_t \tilde{\pi}\left(\omega_t\right) \tag{2}$$

where $\omega_t \equiv w_t/A_t$ is the productivity adjusted skilled wage and $\tilde{\pi}\left(\omega_t\right)$ is productivity adjusted profit.

1.3 Balanced Growth

When defining the amount of skilled labor to employ in the discovery process, a firm maximizes the flow of expected profit from discovering and patenting the invention.

Denote by V_{t+1} the value of generation $t+1$ of inputs: as the present value of a higher quality intermediate good appears a time δ after its discovery, and because any firm which competes in research the term has a probability $e^{-\lambda\delta \int_0^\delta n_t(s)ds}$ to win the R&D race, the flow of profits from research writes:

$$e^{-\lambda\delta \int_0^\delta n_t(s)ds} \lambda V_{t+1} e^{-r\delta} - w_t n_t$$

The integral $\int_0^\delta n_t(s)ds$ mirrors the uncertainty of the innovation process: because the time of the new discovery is unknown, the longer the time from the beginning of R&D race, the lower the success probability from an individual firm's perspective if no invention has been marketed during that time, and the lower the incentive to conduct research, namely the skilled labor to employ in research. It is easy to show that this integral converges to a stationary equilibrium and the equilibrium is asymptotically stable.

Proposition 1 $\int_0^\delta n_t(s)ds$ *has a stationary solution* n_t. *This solution is asymptotically stable.*

Proof. See appendix.

From standard computations we get the following arbitrage condition:

$$w_t = e^{-\lambda\delta n_t} \lambda V_{t+1} e^{-r\delta} \tag{3}$$

Each innovation is assumed to be drastic, namely each type of inputs is replaced by the next vintage. So, while the patent lasts forever, the monopoly rent expires at the *time* when a new innovation occurs in the market (A-H 1992, p.328). This time is postponed for δ compared with the discovery time. Thus, the present value of returns from discovering and patenting the higher quality intermediate goods is equal to the profit flow $\int_0^\infty \pi_{t+1} e^{-r\tau} d\tau$, stemming from the innovation $t+1$, minus the capital loss suffered from being replaced, with probability λn_{t+1}, by a new innovator at time $\tau_{t+1}+\delta$, where τ_{t+1} is the discovery time of the generation $t+1$ of inputs.

From the discovery of new inputs up to the replacement of the older ones, the current innovator continues to earn a flow of profit equal to:

$$\bar{V}_{t+1} = \int_0^\delta \pi_{t+1} e^{-rs} ds \tag{4}$$

Then, the capital loss is given by $\left(V_{t+1} - \bar{V}_{t+1}\right)$ instead of V_{t+1} and the asset equation writes:

$$V_{t+1} = \left[\pi_{t+1} - \lambda n_{t+1} \left(V_{t+1} - \bar{V}_{t+1}\right)\right] / r \tag{5}$$

Notice that the value of capital loss, accounting for the time when the invention is marketable, results lower than the one which would have been suffered if invention would have been immediately patentable.

By combining both (5) and (4) , we finally get:

$$V_{t+1} = \frac{\pi_{t+1} + \lambda n_{t+1} \bar{V}_{t+1}}{r + \lambda n_{t+1}} = \frac{\left[r + \lambda n_{t+1} \left(1 - e^{-r\delta}\right)\right] \pi_{t+1}/r}{r + \lambda n_{t+1}} \tag{6}$$

As usual, in (6) the denominator includes both the interest rate r and the creative-destruction rate λn_{t+1}: the higher the research level in the next period, the higher the creative-destruction rate and the lower the present value of an innovation.

By combining (3) and (6), and given the relationship between $\pi(w)$ and $\tilde{\pi}(\omega)$, the arbitrage condition can be re-written as:

$$\omega_t = e^{-\lambda\delta n} e^{-r\delta} \lambda \frac{\gamma \left[r + \lambda n_{t+1}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}} \tag{7}$$

Our balanced growth analysis follows the one performed by Aghion-Howitt (92). The amount of research employment $\bar{n}$ at a stationary equilibrium is defined as follows:

$$\frac{\omega}{e^{-\lambda\delta\bar{n}} \lambda e^{-r\delta}} = \frac{\gamma \left[r + \lambda\bar{n}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}\left(\omega\left(N - \bar{n}\right)\right)/r}{r + \lambda\bar{n}} \tag{8}$$

$\bar{n}$ is higher the lower the interest rate r as a reduction of r positively affects V, the higher the size of innovation γ as it increases the profits attainable from the new discovery, and the higher the arrival rate of innovation λ as it improves the productivity in research. Finally, the research employment at equilibrium still increases with the labor force L, as it reduces the wage rate of skilled workers. Yet, the discovery's uncertainty stemming from the information-dissemination lag introduces a congestion in the research sector: the higher the amount of labor devoted to research, the lower the chance of winning the R&D race from an individual firm's perspective and the less profitable to invest in R&D activities. Then, the following holds:

Proposition 2 *Any increase in the labor force L, namely in the population size, does not raise the research labor proportionally.*

Proof. See appendix.

As far as balanced growth is concerned, the final output good during the time interval t can be written as

$$y_t = A_t(L - \bar{n})^{\alpha}$$

from which it follows that

$$y_{t+1} = \gamma y_t$$

As in Aghion and Howitt (1992 and 1998) we can write

$$\ln y\,(\tau + 1) = \ln y\,(\tau) + \varepsilon\,(\tau)$$

where ε the number of innovations between τ and $\tau + 1$. Research-firms need the current innovation to be disclosed before thinking of a newer generation of inputs, and an innovation is disclosed only when the patent process ends, then there is a lag between the time when the research process t closes and the one $t+1$ starts. Thus, the number of innovations ε is bounded above. Further, due to the information-dissemination lag δ on the current invention, no firm starts to work on the next vintage, while investing with nil probability of success on the current invention if the discovery has occured although not been disclosed. The research process is governed by a Poisson process, then the average growth rate AGR of the economy turns out to be:

$$g = \sum_{j=0}^{\kappa} \frac{(\lambda n \tau)^j \, e^{-\lambda n \tau_\delta}}{j!} j \ln \gamma = \lambda m \left(n\right) \ln \gamma \qquad (9)$$

where $\lambda m \left(n\right) = \sum_{j=0}^{\kappa} \frac{(\lambda n \tau)^j e^{-\lambda n \tau_\delta}}{j!} j$ which is bounded above by $\kappa \equiv \frac{1}{\delta}$. The lag in disseminating informations eliminates the well-known scale effect which has been proved by Jones (1995) to be at odds with the empirical evidence. The scale effect predicted by several endogenous growth models states that if the amount of resources in the research sector doubles, then the per-capita growth rate should double as well Recently, several models have been developed with the aim of avoiding this prediction. A first class of models (Jones 1995, Kortum 1997, and Segerstom 1998) assumes that the arrival rate of innovation decreases with level of knowledge, namely a larger economy needs to invest a larger amount of resources in R&D to get a constant growth rate of productivity as the number of goods to be improved is higher. An alternative approach (Dinopoulos and Segerstom 1998, Howitt 1999, Peretto 1998, Young, 1998) assumes that research can increase either the quality of goods within a product line or expand the number of the available goods. Then, the growth rate is determined by the research investment in each product line. As increasing the population size is reflected on extending proportionally the numbers of product lines, without affecting the research effort in each sector, the scale effect is totally dissipated. Although these formulations represent advancements in removing the scale effect, their findings are sensitive to the value of parameters on which the analysis is performed. Concerning the first class of models, the scale effect disappears only if diminishing returns in the production of new ideas are imposed. Further, according to this formulation policy does not affect the long-run growth rate, which is in sharp contrast with the endogenous growth theory's flavour. This scale effect problem is not even solved in the alternative approach. Adding a new dimension to the innovation activity, namely horizontal innovation, removes the scale effect only if the number of the product lines grows proportionally with population; otherwise, if it increases less *(res. more)* than proportionally, a positive *(res. negative)* scale effect still occurs. In this slightly modified version of Aghion and Howitt model (1992), the information lag puts an upper bound to the average growth rate such that the scale-effect of Aghion and Howitt (1992) disappears while leaving unaffected the main properties of the original model.

Then, we can claim the following:

Proposition 3 *The information/dissemination lag puts an upper bound to per-capita growth rate, regardless of the amount of skilled labor, thereby eliminating the well known scale effect property proved by Jones (1995) to be inconsistent with existing data.*

1.4 Concentration and Cooperation in the R&D Sector

The main feature of the decentralized setting is the congestion phenomenon due to the information gap affecting negatively research effort. This is rather realistic in an economy in which a large number of small firms undertake independent R&D projects. In the model of the previous section this was stylized by the assumption of zero-measure firms. Though in the real industrial world it is frequent for small firms to exchange some relevant information with each other, the character of a firm as an information barrier (Jovanovic and Rob 1989, Jovanovic and MacDonald 1994, Jovanovic and Rousseau 2001) provides strong incentives for firms to merge their R&D units or to let them cooperate (Cozzi 1999). In this model with free entry, tradable factors, and constant returns, homogeneous R&D workers extract all producers' surplus generated: so large firms can be equivalently seen as skilled workers' R&D associations. The aim of reducing the information dissemination lag provides a natural motivation for R&D firms to merge and/ for R&D workers to cooperate with each other. Of course, the antitrust authorities should worry that a cartel of R&D producers might refrain from innovation by internalizing the business stealing externality. Then we investigate which firm size is the more adequate system to overcome this sort of congestion without discouraging aggressive innovative activity. The aim of this section is precisely to define the number of research-firms that maximizes growth, verifying the relationship between the number of research-firms and research effort. We can assume that the antitrust authorities impose that each R&D firm size be not larger than one N^{th} of the market. Since all researchers would want to work in the biggest R&D firms (because the duplication risk is smaller) exactly N research-firms participate in the R&D race. Each research-firm i devotes exactly n_t^i to R&D sector with

$$n_t^i \equiv \frac{n_t}{N}$$

where n_t is the total amount of labor employed in the research-sector.

In the determination of its R&D effort each firm cares about the probability of winning the R&D race. This is inversely related to R&D effort of the other research-firms.

Focusing on symmetric equilibria we can re-write the arbitrage condition for the firm i as follows:

$$\omega_t = e^{-\lambda \frac{(N-1)}{N} n_t \tau_\delta} e^{-r\delta} \lambda \frac{\gamma \left[r + \lambda n_{t+1} \left(1 - e^{-r\delta} \right) \right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}} \tag{10}$$

where $e^{-\lambda \frac{(N-1)}{N} n_t \tau_\delta}$ refers to the probability of success for the firm i. It is expressed in terms of the number of research-firms working in the market and total amount of labor employed in R&D. It easy to see that RHS is a decreasing function of N. The following holds:

Proposition 4 *The higher the number of non-cooperating research firms, N, the lower the probability of a firm's winning the R&D race, the lower the expected value of an hour in research. The higher the number of firms in the market the stronger the effects of the information-gap, the lower the expected aggregate growth rate, and the lower the skilled/unskilled wage gap.*

Proof. See appendix.

As the R&D effort is inversely related to the number of research-firms and the higher the research-effort the higher the average growth rate of the economy the best solution under laissez-faire is obtained when a minimum number of firms operate in the market that is when $N = 2$. Since skilled workers are always better off the relatively larger their R&D firm we can conclude:

Corollary 5 *An antitrust policy that seeks to maximize growth would allow the formation of a non-collusive duopoly in the R&D sector, which would immediately form.*

Notice that here the duopoly emerges endogenously as R&D workers would tend to unify their effort, but the antitrust authorities can allow or ban it. Hence at the R&D stage every duopolist is under the constant threat of possible free entrants (defector groups of workers). Therefore all profits are dissipated into skilled wage.

1.4.1 Monopoly or Duopoly?

Should the antitrust authority allow R&D firms to join a unique firm, that will become the unique intermediate good monopolist? We just need to compare (10) with $N = 2$ to the following arbitrage condition of the monopolist:

$$\omega_t = \lambda \frac{(\gamma - 1)\,\tilde{\pi}_{t+1}}{r} \tag{11}$$

and conclude that the larger the information lag, the higher the quality jump, and the more populated the economy the more innovative the monopoly compared to the duopoly. Reminding the reader that in this representative sector model the monopoly should be correctly interpreted as belonging to an infimum sub-sector of the whole economy, it may be interesting to note that our analysis implies that - *ceteris paribus* - when a large area integrates economically it is more likely that antitrust authorities concerned with innovation would decide to allow more concentration than in smaller areas. The reason is that with a large skilled labor mass working for the rival each duopolist would seriously fear that its potential discovery could already be - or will soon become - obsolete.

1.5 Endogenous Skills

Up to now we have being following Aghion and Howitt's (1992) assumption of a fixed composition of the labor supply between skilled and unskilled workers. In this section we show that our analysis allows to relate efficiency enhancing mergers to the endogenous composition of the labor force.

Indeed it is straightforward to endogenize the supply of skills in this economy, by adapting Dinopoulos and Segerstrom's (1999) methodology to the closed economy framework . We assume that individuals are finitely lived members of infinitely lived households, being continuously born at rate β, and dying at rate d, with $g = \beta - d > 0$; let L_0 the population size at time zero, then at time τ it results equal to $L(\tau) = L_0 e^{g\tau}$. $D > 0$ denotes the exogenously given duration of their life. Then, as the number of births at time τ is equal to the number of deaths at time $\tau + D$, it follows that $\beta = ge^{gD}/(e^{gD} - 1)$. People are altruistic in that they care about their household's total discounted utility according to the usual intertemporally additive functional. They choose to train and become skilled at the beginning of their lives, and the (positive) duration of their training period - in which the individual cannot work - is exogenously fixed as $T < D$.

Hence an individual with ability $\theta \in [0, 1]$ uniformly distributed decides to train if and only if the following is satisfied:

$$E_\tau \left[\int_\tau^{\tau+D} e^{-r(s-\tau)} w^u(s) ds \right] < E_\tau \left[\int_{\tau+T}^{\tau+D} e^{-r(s-\tau)} \max(\theta - \gamma, 0)\, w(s) ds \right], \tag{12}$$

with $0 < \gamma < 1/2$, and with E_τ denoting expectations as of time τ. Notice that an individual of ability $\theta > \gamma$ is postulated able to accumulate skill (human capital) $\theta - \gamma$ after training, while individuals with too low ability ($\theta < \gamma$) never get any skill from schooling.

As Dinopoulos and Segerstrom (1999) we focus on the steady state (balanced growth) analysis, in which all variables - and real wages in particular - are expected to grow at the same constant expected growth rate as A_t. Hence any individual decides to get an education if and only if her ability level is no less than

$$\theta_0 = \sigma \frac{\omega^u}{\omega} + \gamma,$$

where $\sigma > 0$ is a parameter dependent constant and θ_0 results from (12) holding as an equality. Therefore the aggregate supply of unskilled labor at time τ is

$$M(\tau) \equiv \theta_0 L(\tau) = \left(\sigma \frac{\omega^u}{\omega} + \gamma\right) L(\tau) \tag{13}$$

while the remaining fraction of population $(1-\theta_0)\, L(\tau)$ is either involved in training or skilled. Accounting for the duration of the training process, it is easy to define a subpopulation of individuals born between $\tau - D$ and $\tau - T$ which is skilled at some time τ, namely:

$$\int_{\tau-D}^{\tau-T} (1-\theta_0)\, L(s) ds = (1-\theta_0)\, L(\tau) \left(e^{g(D-T)} - 1\right) / (e^{gD} - 1)$$

or

$$(1-\theta_0)\, L(\tau)\phi$$

where $\phi = \left(e^{g(D-T)} - 1\right) / (e^{gD} - 1)$.

As the average of skilled workers completing the training process is equal to $[(\theta_0 - \gamma)/2 + (1-\gamma)/2]$, then the aggregate supply of skilled labor at time τ in units of efficiency is:

$$L^s(\tau) = (\theta_0 + 1 - 2\gamma)\,(1 - \theta_0)\,\phi L(\tau)/2 \tag{14}$$

with $0 < \phi < 1$.

Since the fraction of the population that decides to get skilled is only dependent on the productivity adjusted skilled/unskilled wage ratio and on parameters the results of the previous sections continue to hold and we can state:

Proposition 6 *The lower the number of non-cooperating R&D firms, N, the higher expected aggregate growth rate, the higher the skilled/unskilled wage gap, and the larger the fraction of population that decides to get skilled.*

Proof. Just repeat the proof of Proposition 1 using (13) and (14) instead of $M = 1$ and L. **Q.E.D.**

1.6 Conclusion

In this paper we have shown that due to the information gap between research firms a kind of congestion appears in the R&D sector. This phenomenon, reducing the research effort, affects growth negatively. Moreover as R&D effort is inversely related to the number of non-cooperating research firms, we conclude that the best growth-enhancing policy under *laissez-faire* is to allow at least a contestable duopoly for each innovative industry. It seems interesting that the information dissemination lag allows us to capture in a unified framework phenomena — such as wage distribution, R&D cooperation and their interplay — analyzed in often unrelated investigations, although with similar time pattern.

We can re-interpret some recent trends in the US and EU in the light of our simple model's results. Since the mid 70's - also due to weaker international barriers and fierce competition by newly industrialized countries - US and EU authorities have relaxed their antitrust policy against either mergers or R&D co-operation in order to improve their competitiveness. This has stimulated R&D efforts, reduced innovation delay (see Agarwal and Gort, 2000), improved product qualities at a faster rate, and at the same time it has increased wage inequality and the fraction of the highly educated population.

Our results seem consistent with some well known empirical evidence on intefirm barriers, the related incentives for firms to cooperate, and the plausible link between market structure, wage inequality, and growth. Neverthless, they have been obtained within the extremely stylized Aghion and Howitt's (1992) framework. It would be interesting to extend our logic to an explicitly more complex framework, such as for example Aghion et al. (2001). Moreover, by analyzing US data in the period 1985-1998 Link, Paton and Siegel (2002) find that authorities systematically allow R&D agreements more often when the country is losing international competitiveness in high technology sectors. This finding is consistent with our claim that policy changes are at the root of the observed consequence for innovation and income distribution. At the same time it suggests to endogenize industrial policy in an explicit international setting. These will be topics for further work.

1.7 Appendix

Proof of Proposition 1. Let $\omega(x) \equiv F'(x) + xF''(x)$ be the marginal-revenue function. As $F(x) = Ax^{\alpha}$, then the marginal-revenue function can be manipulated and expressed as follows:

$$\omega(x) = \alpha^2 x^{\alpha-1}$$

Combining the previous expression with (7) and (1), we get:

$$\alpha^2(L_t - x_t)^{\alpha-1} = e^{-\lambda\tau_\delta \int_0^\delta n_t(s)ds} e^{-r\delta} \lambda \frac{\gamma \left[r + \lambda n_{t+1}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}}$$

or

$$\dot{Z} = L - e^{\lambda\tau_\delta/(1-\alpha)Z}(W/\alpha^2)^{1/(\alpha-1)} \tag{15}$$

where

$$W = e^{-r\delta}\lambda \frac{\gamma \left[r + \lambda n_{t+1}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}}$$

and $Z = \int_0^\delta n_t(s)ds$, given $Z(0) = 0$.

Standard computations yields to

$$\bar{Z} = \ln(L/\tilde{W})/(\lambda/(1-\alpha))$$

where $\tilde{W} = (W/\alpha^2)^{1/(\alpha-1)}$.

As it follows that $F'(\bar{Z}) < 0$ for any value of parameters, $\bar{Z}$ turns out to be an asymptotically stable rest point. **Q.E.D.**

Proof of Proposition 2. By manipulating the (8) it follows

$$\omega = e^{-\lambda\delta\bar{n}}\lambda\gamma e^{-r\delta}\frac{\left[r+\lambda\bar{n}\left(1-e^{-r\delta}\right)\right]\left(\frac{1-\alpha}{\alpha}\right)\omega x/r}{r+\lambda\bar{n}} \tag{16}$$

and hence

$$\frac{e^{\lambda\bar{n}\delta}e^{r\delta}}{\lambda\gamma\bar{n}} = \left(1-\frac{e^{-r\delta}}{\frac{r}{\lambda\bar{n}}+1}\right)\left(\frac{1-\alpha}{\alpha}\right)\left(\frac{L}{\bar{n}}-1\right)/r$$

from which easily we get

$$\frac{\vartheta\left(L/\bar{n}\right)}{\vartheta L} > 0$$

Q.E.D.

Proof of Proposition 3. Let us rewrite the arbitrage condition as

$$\phi\left(n_t, N\right) \equiv e^{-\lambda\frac{(N-1)}{N}n_t\delta}e^{-r\delta}\lambda\frac{\gamma\left[r+\lambda n_{t+1}\left(1-e^{-r\delta}\right)\right]\tilde{\pi}_{t+1}/r}{r+\lambda n_{t+1}} - \omega_t = 0$$

Then according to the implicit function theorem we know that

$$\frac{\partial n_t}{\partial N} = -\frac{\partial\phi\left(\bullet\right)/\partial N}{\partial\phi\left(\bullet\right)/\partial n_t}$$

where

$$\frac{\partial \phi\left(\bullet\right)}{\partial N} = -\lambda n_t \delta N^{-2} e^{-\lambda \frac{(N-1)}{N} n_t \delta} e^{-r\delta} \frac{\gamma\left[r + \lambda n_{t+1}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}}$$

and

$$\frac{\partial \phi\left(\bullet\right)}{\partial n_t} = -\lambda \frac{(N-1)}{N} \delta e^{-\lambda \frac{(N-1)}{N} n_t \delta} e^{-r\delta} \frac{\gamma\left[r + \lambda n_{t+1}\left(1 - e^{-r\delta}\right)\right] \tilde{\pi}_{t+1}/r}{r + \lambda n_{t+1}}$$

So it follows

$$\frac{\partial n_t}{\partial N} = -\frac{\partial \phi\left(\bullet\right)/\partial N}{\partial \phi\left(\bullet\right)/\partial n_t} < 0$$

since $\frac{\partial \phi(\bullet)}{\partial N} < 0$ and $\frac{\partial \phi(\bullet)}{\partial n_t} < 0$ that is the lower the number of research-firms the higher the equilibrium effort in R&D and the higher the average growth rate of the economy. The other statements follow straightforwardly. **Q.E.D.**

CHAPTER 2.

Capacity Expansion and Dynamic Monopoly Pricing

S. Demichelis and O. Tarola

Abstract.

We substitute to the plant size problem, as investigated by Chenery (1952), a new version in which a profit-maximising monopolist may combine its investment policy with a price policy adjusting demand upwards or downwards over time. We characterize the optimal price and investment policies. The optimal price policy determines an investment pattern either with constant increments of capacity over time, or becoming constant after a finite time. The existing capacity is either fully used at each instant between two investment dates; or the monopolist first quotes the instantaneous monopoly price and, thereafter, the price dampening instantaneous demand at the optimal installed capacity level

2.1 Introduction

The so-called plant-size problem introduced in the fifties by Chenery (1952) studies the optimal capacity to install in order to serve an exogenous demand increasing over time, with economies of scale in plant construction. This problem arose in the context of developing countries: explosive demographics forced the governments to regularly increase capacity over time, in order to face increase in demand resulting from such populations' booms. Increasing returns to scale in plant construction opened the door to the possibility of anticipating the increments of capacity required to meet future demand levels, and exploiting thereby the resulting economies of scale. In this context, it was reasonable to assume that demand expansion was pushed exogenously by the population growth.

Today, in developed countries, similar demand booms are still observed in several industries, but for other reasons than an exogenous population growth. These booms now often follow from the globalisation of trade: while markets in developed countries expanded initially around local or national demands, they are now progressively concerned by larger and larger geographical areas. In some sense, firms operating in these industries are faced today in developed countries with a problem akin to the problem met by governments in developing countries in the past: they also expect progressive expansion of demand consecutive to trade globalisation,while benefiting as well from economies of scale in plant construction. However, in this new context, it is not reasonable to view demand increases addressed to firms operating in such industries as exogenous. These firms are no longer public institutions, as in the original situation, but rather profit maximizing institutions. As such, they benefit from specific instruments which were not at the disposal of the governments. In particular, they can now accompany their investment policy by a price policy through which they can manipulate the levels of demand through time.

In order to capture this new strategic possibility, we study in the present paper a similar plant size problem as in Chenery (1952) and Manne (1961). Yet we now allow the profit-maximising firm producing the good to combine its investment policy with a product price policy adjusting demand upwards, or downwards, over time. This extension has to cope with the difficulty that, even if the volume demanded is exogenously increasing (or decreasing) at each price over time, the firm can now decrease or increase it by instantaneous increases or decreases in the price pattern. We characterize the optimal price and investment policies of the monopolist under the assumption that the firm while controlling the size of the investments to be undertaken, does not control the dates, which are assumed to be equally spaced through time, at which these investments have to be consented. Our findings are as follows. The optimal price policy, through which the monopolist can possibly dampen or enhance the expansion of demand through time, leads to an investment pattern with increments of capacity which are either constant over time at each investment date, or start to become constant after a finite set of investment dates. Furthermore, the optimal price pattern between two dates at which new capacity is installed falls into two categories: either the price pattern is such that total existing capacity is fully used at each instant between these dates; or the instantaneous monopoly price is used for some period within the cycle and, thereafter, until the end of the cycle, the price which dampens instantaneous demand at the level of installed capacity is adopted. Finally, we show that the optimal constant increment of capacity is smaller than the one which would have been selected by a planner facing a demand growing exogenously at a rate corresponding to the instantaneous monopoly price, as in Manne (1961). These conclusions are *a priori* far from being evident. Even constrained to invest at equally spaced dates, the monopolist could have as well preferred to manipulate the price, according to the value of the interest rate, in order to decrease or increase the size of the investments through time rather than using a constant–cycle investment size policy.

Research on the plant-size problem has been pursued without discontinuity since the sixties. Manne (1961) and (1967) faced the capacity expansion problem which was met by several manufacturing facilities in India. Assuming an exogenous demand growing linearly over time and economies of scale in investment costs, he finds that the optimal investment policy is constant-cycle : successive investments are all of the same size and undertaken at equally spaced points of time. These findings have been extended afterwards in several directions. Srinivasan (1967) proves that the constant-cycle property of the optimal plan extends to the case where demand is growing geometrically over time. Gabszewicz and Vial (1972) extend the analysis to exogenous technological progress which can arise either at a deterministic date or at a random time in the future. In both cases, the constant-cycle property of the optimal policy still holds. Around the same period a considerable amount of research work has been devoted to the problems of lead times and uncertain demand (see in particular Nickell (1977), Freidenfelds (1981), Bean and al. (1992), Chaouch and Buzacott (1994) and Ryan (2002)).

Although these research contributions represent important advances in the understanding of the plant-size problem, they all assume that demand for capacity increases is exogenously given, and cannot be adjusted through some price policy. Of course, this assumption was a natural entry point since past contributions were mainly concerned with finding the solution of a planning problem met in the framework of a country development program, and not with identifying the solution of a profit-maximiser monopolist. Nevertheless, as shown in the present paper, the plant-size problem can be reformulated in order to take into account this market alternative interpretation, in which the firm is allowed to manipulate instantaneous demand through price. So, departing from the above research lines, we propose to relax the assumption of exogenous demand increase. We replace it by an assumption about the time expansion of a price-quantity relation specifying how instantaneous demand varies with price at each point of time. More precisely, we suppose that, at each point of time, market demand is given by a linear function of instantaneous price, the intercept of which increases linearly with time. Moreover, we make the assumption alluded above concerning the size of the investments, namely, we fix a sequence of equally spaced points of time at which the investments for adding capacity may be undertaken and characterise the optimal policy of the monopolist in terms of his instruments: price regime and investment's size.

The model is described in section 2. In the same section, we fully characterise both the optimal price and investment policies .We summarize our findings in the conclusion and propose some paths for further research.

2.2 The analysis of the optimal policy

2.2.1 The model

We consider a monopolist facing a demand function $D(t,p)$ defined by

$$D(t,p) = At - p$$

with t denoting continuous time and p instantaneous price: demand at each price is accordingly expanding through time because the intercept of the demand function increases at a rate proportional to t. At each instant of time t, the capacity of the firm is bounded by the existing amount of equipment, which we denote by $X(t)$. While the existing capacity may exceed the current demand level $D(t,p)$, no undercapacity is admitted, so that $D(t,p) \leq X(t)$. The investment cost for adding new capacity x, which includes both fixed cost k, $k > 0$, and variable cost ax, exhibits economies of scale, namely

$$f(x) = k + ax;$$

This cost structure is assumed to hold for ever (no technological progress). Also we assume no capital depreciation. Time is discounted at a constant interest rate r. The time horizon is unbounded.

Now we assume that the monopolist is allowed to make investment decisions x_{t_i} at fixed equally spaced points of time t_i, $i = 0, 1, 2, ...$ In the interval $[t_i, t_{i+1}[$,the existing capacity remains constant and is equal to $X(t_i) = \sum_{k:t_k \leq t_i} x_{t_k}$. At each t_i, the firm decides the increment of capacity x_{t_i} to be installed and sets the price $p(t)$ for $t \in [t_i, t_{i+1}[$. Then at time t_{i+1} a new investment is undertaken and a new price schedule for the period is defined, and so on. Up to a change of units we set $t_i = i$. The interval of time $[i, i+1[$ between two dates at which a new investment is decided is called a *cycle.*

Formally, given the sequence of investment dates $\{i\}$, $i = 0, 1, 2, 3...$,the problem of identifying the *optimal policy* for the monopolist consists in finding $\mathbf{x} = (x_0; x_1; ....x_i...) \in R_+^\infty$ and $p(t)$, measurable function of $t \in R_+$ so that the objective function $V(\mathbf{x}, p(t))$defined by

$$V(\mathbf{x}, p(t)) = \int_0^\infty p(t) D(t, p(t)) e^{-rt} dt - \sum_{i=1}^\infty (k + ax_i) e^{-ri}$$

is maximized, subject to the capacity constraint

$$D(t, p(t)) \leq X(t).$$

The definition of optimality decomposes into two components. The first refers to the properties of the optimal price pattern trough time while the second concerns the optimal sequence of capacity increments required in order to meet demand at each instant of time.

2.2.2 The optimal price policy within a cycle

In order to find the optimal price solution we first remark that the cost function $\sum_{i=0}^{\infty}(k+ax_i)$ does not depend on the choice of $p(t)$. Accordingly, a sufficient condition for the optimality of $p(t)$ is that it maximizes the integrand $p(t)\,D(t;p(t))$ at any point t, given the capacity constraint.

Then we observe that, if the investment size decision $\mathbf{x}$ is kept fixed from one period to the other, $V(\mathbf{x},p(t))$ achieves its maximum for $p(t)$ given by

$$p(t)=\max(\frac{At}{2},At-X_i)$$

for $i=[t]$ the integer part of t. This follows simply from the maximization problem

$$\begin{gathered}\underset{p(t)}{max}=p(t)D(t;p(t))\\ \text{s.t. }\ D(t;p(t))\leq X_i, i=[t]\,.\end{gathered}$$

During the cycle $[i,i+1[$, there are two possible price regimes depending on the capacity level compared with the demand level. Assume that at some date t, $i\leq t<i+1$, the capacity constraint is not binding, namely $D(t;p(t))<X_i$. Then, at the optimal policy, $p(t)$ should be set equal to the maximizing price $p^M(t)=\frac{At}{2}$ as the demand does not need to be dampened. Yet, the demand expands over time while the current capacity remains fixed. When it happens that the capacity constraint turns out to be binding, namely $D(t;p(t))=X_i$, then the firm has to choose the price $p^C(t)=At-X_i$ so as to contract the demand $D(t;p(t))$ within the limits imposed by its plant size. These two price patterns $p^M(t)$ and $p^C(t)$ are called, respectively, *monopoly price regime* and *constrained price regime,* and the resulting demand patterns are denoted by $D^M(t)$ and $D^C(t)$, respectively.

Further, we denote by t_i^* the point of time when the monopolist switches from the monopoly price regime into the constrained regime (namely when $p^M(t)$ becomes equal to $p^C(t)$); we label this point as the *switching point* t_i^*. It is easy to see that $t_i^* = \frac{2X_i}{A}$. Then, four scenarios may arise. Assume first that the switching point lies after $i+1$ or that it is exactly equal to $i+1$; then the optimal price pattern coincides with the monopoly price regime during the whole cycle (scenarios A and B respectively). Assume now that the switching point lies between the two investment dates of a cycle $[i, i+1[$; then both these two regimes are used at the optimal price pattern, the first one between i and t_i^* and the second between t_i^* and $i+1$ (scenario C). Assume now that the switching point lies before i; then the monopolist is forced to use the constrained price regime during the whole cycle in order to meet the capacity constraint (scenario D). Figures 1 and 2 illustrate these 4 scenarios.

It is easy to see that this switching point t_i^* can never exceed the point $i+1$: in this case, it would be sufficient to reduce the installed capacity up to the point where this capacity is equal to $D^M\left(i+1, p^M(i+1)\right)$, say by ε, in order to gain the discounted costs $\varepsilon a\left(e^{-ri} - e^{-r(i+1)}\right)$ without incurring any reduction of revenue. So the first scenario never arises. We prove now that the switching point t_i^* can never be *exactly equal* to the regeneration point $i+1$, excluding thereby also scenario B.

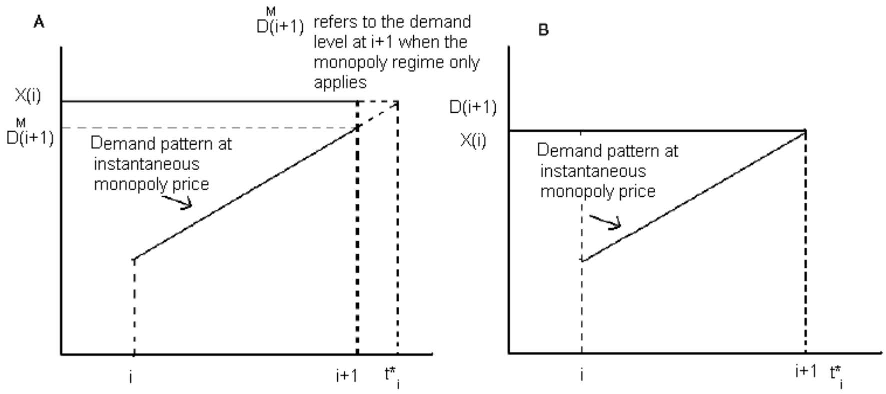

Figure 1: The switching point t_i^* lies either exceeds (scenario A) or coincides with the point of time $i+1$ (scenario B).

Lemma 7 *During any cycle i, the switching point t_i^* belongs either to the interior of the cycle i, or it is strictly smaller than i.*

Proof. Assume that, for some i, $X_i^{'}$ would be the optimal installed capacity at date i and that $X_i^{'} = D^M\left(i+1\right)$. Then the monopolist can use the monopoly regime $p_i^M(t)$ during the whole cycle. The present value of the discounted flow of revenues R_i during the cycle i obtains as

$$R_i = \int_i^{i.+1} \left(\frac{At}{2}\right)^2 e^{-rt} dt. \tag{17}$$

Now assume that the installed capacity $X_i^{'}$ drops by a small quantity ε, so that $X_i^{'} - \varepsilon < D^M\left(i+1\right)$ with $0 < \varepsilon < 1$. The monopolist gains the discounted cost saved by reducing the investment by ε,namely $\varepsilon a\left(e^{-ri} - e^{-r(i+1)}\right)$. Yet, the demand is not completely met. This induces to switch from the monopoly price regime p_i^M to the constrained price p_i^C. Thus, the present value of the discounted flow of revenues during the cycle i turns into

$$\begin{aligned} R_i' &= \int_i^{i+1-\delta} \left(\frac{At}{2}\right)^2 e^{-rt} dt + \qquad (18) \\ & \int_{i+1-\delta}^{i+1} \left[\left(\frac{A(i+1-\delta)}{2}\right)^2 + \frac{A}{2}^2 (i+1-\delta)(t-(i+1-\delta))\right] e^{-rt} dt, \end{aligned}$$

where $\delta = \frac{A\varepsilon}{2}$ and the second term corresponds to the revenue stemming from the use of the constrained regime between $i+1-\delta$ and $i+1$. Substracting (18) from (17) , the loss derived from the switch between the two price regimes writes as

$$L = \int_{i+1-\delta}^{i+1} \left[\left(\frac{At}{2}\right)^2 - \left(\frac{A(i+1-\delta)}{2}\right)^2 - \frac{A}{2}^2 (i+1-\delta)(t-(i+1-\delta))\right] e^{-rt} dt.$$

The loss L is of third order in ε, as it is given by the integral of a function of the order of δ^2 over an interval of length δ. The gain $G = \varepsilon a \left(e^{-ri} - e^{-r(i+1)}\right)$ is of first order in ε. Accordingly, for ε small enough, the net loss should be negative, which is the desired contradiction. **Q.E.D.**

We deduce immediately from the above that only the two remaining scenarios can be observed at an optimal price policy. Thus, the optimal price pattern within a cycle must either consist of quoting always the instantaneous constrained price, or using first this price and then, after the switching point, the constrained regime. The first alternative should necessarily hold when the investment at time i is so low that it is even not sufficient to meet the monopoly demand $D\left(i, p^M(i)\right)$ at that time.

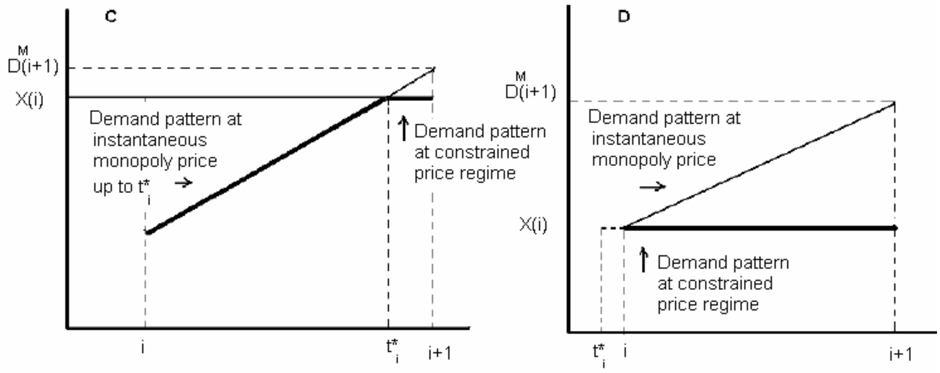

Figure 2: The switching point lies either between i and $i+1$ (scenario C) or before the point of time $i+1$ (scenario D).

The second one corresponds to a situation where the capacity installed at time i is large enough to serve the monopoly demand for some period in the beginning of cycle i, so that the switching point is interior to the cycle. Accordingly , the nature of the optimal price policy within a cycle is directly related to the size of the investment consented at the beginning of the cycle.

2.2.3 The Optimal Investment Policy

Thus, it remains to identify the optimal sequence of investments in order to fully characterize the optimal policy. Clearly, the optimal policy must depend on the main parameters of the model, namely, the interest rate r, the intercepts At of the demand functions and the unit variable cost a. When the unit variable investment cost a is high compared with the growth parameter of demand A, then the revenue derived from deciding a non-negative increment of capacity at date i may well not compensate the increase in investment cost that it entails: in that case, no increment of capacity is consented at the beginning of the cycle so that only the constrained price regime can be adopted within the rest of the cycle, as in figure 2 D. However, as we shall see in the following, this scenario cannot last forever at the optimal investment policy: after a while the increase in demand makes profitable a positive increment of capacity and both the price regimes are then alternated. On the contrary, when the unit variable investment cost a is low compared with the growth parameter of demand A, then the revenue derived from deciding a strictly positive increment of capacity at date i compensates the increase in investment cost that it entails: then at optimum, the price policy alternates instantaneous monopoly and constrained prices, as in figure 2 C. In this case, we prove below that this optimal price policy remains identical in each cycle for ever. The two above cases displayed in figure 2C and 2D, and the ensuing optimal investment policies, are now considered successively. Let us remark the following which will be used in the proof of the next propositions. Consider an optimal $\mathbf{x}^*$ and denote by $V(\mathbf{x})_{i,h}$ the value of the objective function at the vector $(\mathbf{x})_{i,h} = (x_0, ..., x_i - h, ...x_i + h, ..., x_{i+1...})$ for any i and $h \leq x_i$. Clearly it follows from the optimality property of $\mathbf{x}^*$ that $V(\mathbf{x}^*) \geqslant V(\mathbf{x})_{i,h}$ for $(\mathbf{x})_{i,h} = (x_0, ..., x_i - h, ...x_i + h, ..., x_{i+1...})$ for any i and $h \leq x_i$. Then $\frac{d}{dh}V(\mathbf{x})_{i,h} = 0$, $i = 0, 1, 2$, namely the marginal revenue and the marginal costs derived from decreasing by h the capacity installed at time i and increasing by the same amount the one installed at time $i+1$ counterbalance ecah other, for any cycle $[i, i+1[$.

Let us start with the first one and consider the following condition

$$a \geq A\left[1 - \frac{r}{e^{r(i+1)} - e^{ri}}\right] \tag{19}$$

on the parameters. It is easy to see that condition (19) can only hold for a finite number of cycles $[i, i+1[$. Let i_0 be the first cycle at which this inequality is reversed. We have

Proposition 8 *When* (19) *holds, the optimal policy consists, up to the cycle* $[i_0 - 1, i_0[$, *in zero increments of capacity and for* $i \geq i_0$ *in a constant increment of capacity at each cycle. Moreover, the optimal price policy displays first a constrained price regime within the whole cycle, adjusted in each cycle, and then a monopoly and a constrained price regime within each cycle.*

Proof. We consider here the case when the switching point t_i^* is outside of the interior of the cycle $[i, i+1[$. Then it is on the left of i (see figure 2 D), and the installed capacity is not sufficient to meet the instantaneous monopoly demand at i. The constrained price regime only applies. For the F.O.C. to be satisfied at the optimal solution x_i^* for any i, the marginal revenue must be equal to its marginal costs, within any cycle $[i, i+1[$. Then, the following holds:

$$\int_i^{i+1} (At - 2X_i)e^{-rt}dt = ae^{-ri}(1 - e^{-r})$$

with $X_i = X_{i-1} + \lambda_i$. Then, the above condition can be re-expressed as follows

$$\int_0^1 (At - 2X_{i-1})e^{-rs}ds - a(1 - e^{-r}) = 2\lambda_i \int_0^1 e^{-rs}ds \qquad (20)$$

where $s = t - i$ and λ_i is the solution of (20) which does not depend on i. Then,

$$x_i^* = \max[0, \lambda_i]$$

Indeed, for

$$i < i_0,\ x_i^* = 0$$

$$i \geqslant i_0,\ x_i^* > 0$$

where i_0 is the first decision point i such that the revenue $\int_i^{i+1} Ate^{-r(s-i)}ds$ and the cost $a(1-e^{-r})$ resulting from a non-negative investment size exactly counterbalance each other. So, as $x_i^* = 0$ the constrained price regime is operated till i_0. Furthermore, when $x_i^* > 0$ the constrained price regime tends to decrease from cycle to cycle as the investment for adding capacity increases. At the time point when the plant size is so big as to satisfy the demand growing linearly over time, the constrained price regime is replaced by the monopolistic price regime and then we fall in the case when both price regimes apply forever. **Q.E.D.**

Notice that even if (19) holds, in the long run the optimal investment policy is constant cycle and the optimal price policy always consists in alternating monopoly price and constrained price regimes within a cycle. By contrast, as stated in the next proposition the same optimal policies turn out to be observed from the very beginning when the reverse of (19) holds.

Proposition 9 *When the reverse of* (19) *holds, the investment optimal policy is unique and stationary through all cycles.*

Proof. First notice that the switching point t_i^* exists and it is unique for each cycle $[i, i+1[$. Accordingly, if we prove that the optimal investment policy is constant cycle, the optimal price policy repeats identically from a cycle to the other. Consider a specific cycle $[i, i+1[$ and assume that the switching point t_i^* is now in the interior of the cycle. Then, the first order conditions become

$$\int_{t_i^*}^{i+1} (At - 2X_i)e^{-rt}dt = ae^{-ri}(1 - e^{-r})$$

or

$$Ae^{-ri} \int_{t_i^*-i}^{1} (s - (t_i^* - i))e^{-rs}ds = ae^{-ri}(1 - e^{-r}) \tag{21}$$

where it is still $s = t - i$. Then (21) can be rewritten

$$f(t_i^* - i) = a \tag{22}$$

where $f(l) = A \int_{l}^{1} \frac{(s-l)e^{-rs}ds}{(1-e^{-r})}$ is a function that does not depend on i. It is immediate to see that $t_i^* = i + \bar{l}$ where $\bar{l}$ is the unique solution to $f(l) = a$. Then, given the time i when the increment of capacity is installed, t_i^* is univocally determined by $\bar{l}$, which does not depend on i, for any i. Finally, as t_i^* identifies the level of available capacity X_i in any cycle $[i, i+1[$, the solution is stationary, as claimed. **Q.E.D.**

A simple byproduct of the above results is that the optimal constant increment of capacity is smaller than the one which would have been selected by a planner facing a demand growing exogenously at a rate corresponding to the instantaneous monopoly price. If the monopolist would quote at each instant of time the monopoly price, it would generate a linear demand trajectory with slope $\frac{A}{2}$ in the demand-time space (see figure 3). We can find out what is the optimal solution of the planning problem corresponding to this expansion of demand, viewed as in Manne (1961). To this solution, there corresponds an optimal constant cycle in investment and, accordingly, a sequence of equally-spaced dates (*regeneration points* in the language of Manne) at which this constant increment of capacity is consented. Starting with this sequence of cycles, we can wonder whether the constant-cycle optimal investment policy of the monopolist leads at the beginning of each cycle to a capacity increment which is larger or smaller than the one resulting from the planning solution provided by Manne.

Proposition 10 *The optimal constant increment of capacity chosen by the monopolist is smaller than the one which would have been selected by a planner facing demand levels* $D\left(i+1, p^M\left(i+1\right)\right)$ *growing exogenously at a rate corresponding to the instantaneous monopoly price.*

Proof. The demand levels resulting from the use of the monopoly price within the whole cycle, for each cycle, correspond to the price policy in scenario A. In lemma 1, it is shown that this cannot be a price policy corresponding to the optimal pattern selected by the monopolist. Similarly, we have seen that an optimal increment of capacity decided at i cannot strictly exceed the demand level $D^M\left(i+1, p^M(i+1)\right)$. Consequently, the optimal increment of capacity must be strictly smaller than $D^M\left(i+1, p^M(i+1)\right)$ while, at the Manne solution, it is exactly equal to this magnitude. **Q.E.D.**

Thus, whether proposition 8 or 9 applies depends on condition (19), which itself depends on the values of parameters A, a and r. Assume that, at $i=1$, condition (19) holds. Then we know that proposition 8 applies: at the beginning,

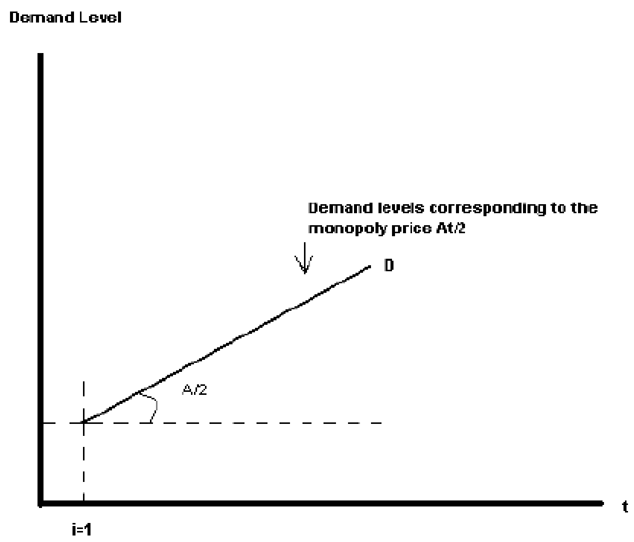

Figure 3: Demand levels corresponding to the monopoly price regime.

for low values of i, no increment of capacity is profitable. However, as the demand function increases over time, at cycle i_0, the increment of capacity starts to be positive and we enter in the scenario described in proposition 9. Notice that, *ceteris paribus,* condition (19) cannot hold for high values of the demand growth rate A. The reason is that, whatever the cost of investing, as expressed by the unit variable cost a and the interest rate r , a sufficiently large value of A should necessarily reverse the inequality (19). Assume now that, at $i = 0$, the reverse of condition (19) holds. Then, from the very beginning, a positive increment of capacity is built, and proposition 9 applies. In the proof of this proposition, we have defined the function $f(l)$ through which the switching point t_i^* of any cycle $[i, i+1[$ corresponding to the optimal policy is derived as

$$f(t_i^* - i) = a$$

with $f(l)$ defined by (22)

Now consider figure 4 where the function $f(l)$ is depicted.

Then, the switching point t_i^* obtained as $i + \bar{l}$, as defined in the proof of proposition 3, is easily identified through figure 4. Moreover, it is immediate that, *ceteris paribus*, when the unit variable cost a increases, the value of $\bar{l}$ decreases; this entails a reduction in the value of the switching point and, accordingly, in the size of the optimal increment of capacity. Notice that when A increases, the whole function $f(l)$ shifts upwards displacing the point $\bar{l}$ to the right. Thus the switching point moves to the right, entailing an increase in the optimal capacity installed at time i. Finally, when the interest rate r increases, it is easy to see that $\bar{l}$ moves to the left, reflecting thereby the increase in investment cost, and entailing the ensuing reduction of installed capacity at the optimum.

2.3 Conclusion

The plant size problem introduced in the fifties was designed in order to solve planning problems raised by an exogenous expansion of demand due to the demographic trend in developing countries. In this paper we have considered an alternative version of the plant size problem, now formulated in the context of a market environment. A monopolist, facing demand expansion and increasing returns, has the opportunity to combine his investment policy with a price policy aiming at manipulating the instantaneous price in order to dampen demand whenever the existing capacity is in a bottleneck. We have characterised both the optimal policies under the assumption that increments of capacity can take place at the beginning of each cycle of a sequence of equally-spaced cycles.

This analysis calls for several generalizations. First, it would be interesting to relax the assumption according to which decision points are fixed, and to analyze the reverse problem, namely, the one when the investment size is fixed and the optimal plan is defined on the basis of both price regimes and investment timing (see footnote 1). Although this would change the frame, yet it would probably not change the main properties of the model. The stationarity property of investment plans involving two price regimes should still hold. Second, it would be natural to introduce other dynamic elements which could influence the optimal time trajectories of prices and investments, like obsolescence (see d'Aspremont *et alii* (1972)) or technical progress, as in Gabszewicz and Vial (1972). Finally, the extension of the model to a dynamic game with two oligopolists would enrich the market

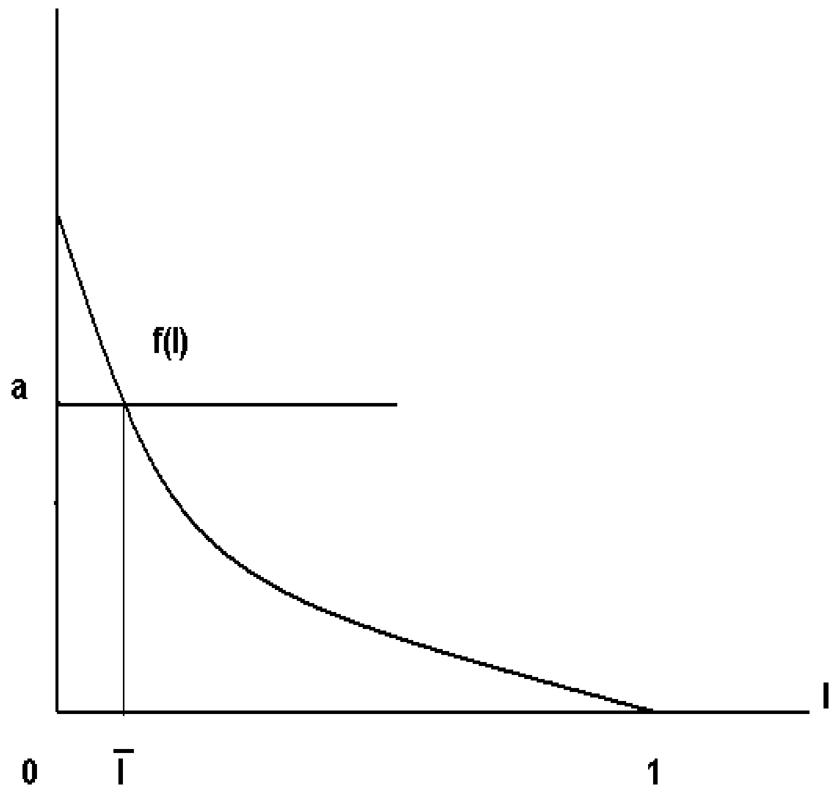

Figure 4: $f(l)$ does not depend on i.

environment in which it could be applied. A natural question in this extended framework would then be: is a similar cyclical behaviour reproduced at the Nash equilibrium of the resulting game? In the longer run, the prospective analysis could even be more ambitious. Since the main outcome of the present paper is to stress the optimal property of successive periods of capacity expansion followed by periods of no investment and contained demand, perhaps a similar model could be used to explain, at least partially, the existence of business cycles (See in particular Abel and Eberly (1996), Caballero and Engel (1991) and (1999)).

2.4 Appendix

2.4.1 Appendix A

The term $\int_{i+1-\delta}^{i+1}\left[\left(\frac{A(i+1-\delta)}{2}\right)^2+\frac{A}{2}^2(i+1-\delta)(t-(i+1-\delta))\right]e^{-rt}dt$ denotes the discounted flow of revenues stemming from the use of the constrained regime during the interval $[i+1-\delta,1+1[$. This can be computed as follows. Let

$$R_L=\int_{i+1-\delta}^{i+1}((At-X_i)X_i)\,e^{-rt}dt \tag{23}$$

be the be the discounted revenue when the monopolist uses the constrained regime within the period $[i+1-\delta,1+1[$. Also, we know that $X_i=\frac{A(i+1-\delta)}{2}$. Then by replacing this value of X_i in (23), we easily get

$$R_L=\int_{i+1-\delta}^{i+1}\left((At-\left(\frac{A(i+1-\delta)}{2}\right)\left(\frac{A(i+1-\delta)}{2}\right)\right)e^{-rt}dt$$

From algebraic manipulations it derives:

$$\int_{i+1-\delta}^{i+1}\left[\left(\frac{A(i+1-\delta)}{2}\right)^2+\frac{A}{2}^2(i+1-\delta)(t-(i+1-\delta))\right]e^{-rt}dt.$$

2.4.2 Appendix B

As stated in proposition 8 and proposition 9, the optimal investment policy can display two different price patterns: either at the beginning of a cycle $[i, i+1[$ the monopoly price regime is adopted and then from the switching point t_i^* up to the end of the cycle the constrained price regime is quoted, or the constrained price regime applies for the whole cycle. Namely when the optimal increment of capacity x_i^* installed at time i suffices to meet levels of demand within the interval$[i, t_i^*]$, the firm charges both the price regimes, otherwise the constrained regime only arises. Without loss of generality, let us consider the first cycle $[0, 1[$ when the plant has still to be initiated. Then, for an initial investment to be non-negative at the optimal solution, the marginal revenue stemming from a non-negative investment must be equal to the marginal cost of installing this capacity, namely:

$$\int_0^1 Ate^{-rt}dt = a(1 - e^{-r})$$

which can be easily solved. So, it writes:

$$\frac{A}{r^2}\left(1 - \frac{r}{e^r - 1}\right) = a$$

Then, when the cost of investing is high with respect to the gain of installing a non-negative capacity or

$$\frac{A}{r^2}\left(1 - \frac{r}{e^r - 1}\right) < a$$

the monopolist refrains from investing, namely $x^* = 0$, and constrained regime only applies; whereas when the cost of investing is low with respect to the gain of installing a non-negative capacity or

$$\frac{A}{r^2}\left(1-\frac{r}{e^r-1}\right)>a$$

then both the regimes apply. Finally, when the growth of demand and investment costs exactly countervail each other or

$$\frac{A}{r^2}\left(1-\frac{r}{e^r-1}\right)=a$$

the firm starts to undertake a non-null investment and the price policy tends to decompose in both the price regimes.

CHAPTER 3.

Monopoly Pricing over Time and the Timing of Investments

Ornella Tarola

Abstract.

We consider a monopoly producing a good whose intercept of its demand function grows linearly over time: at each price demand expands through time as the intercept of the demand function grows. Whatever the price policy that is adopted, the monopolist invests in order to meet the resulting demand growth. The monopolist can only invest in indivisible factory units. We identify the optimal price policy, and the ensuing optimal sequence of investment time points the monopoly selects through time. We show that this sequence satisfies the constant cycle property observed under capacity expansion for an exogenous increase in demand (Manne, 1961).

3.1 Introduction

Adding facilities to optimally meet rising demand for these facilities is one of the most debated issue in all the industries which are characterized by trade-offs in decision making process due to economies of scale in investment costs or commitment of substantial resources for further capacity. Typical examples are given by heavy process industries, communication network, and water resources systems. Research on this topic have been pursued since the sixties in the context of developing countries where explosive demographics forced the governments to regularly increase capacity over time, in order to face increases in demand resulting from such populations' booms. Manne (1961 and 1967) examined the capacity expansion problem which was met by several manufacturing facilities in India. Assuming an exogenous demand growing linearly over time and economies of scale in investment costs, he found that the optimal investment policy is constant-cycle: successive investments are all of the same size and undertaken at equally spaced points of time. These findings were extended in several directions. Srinivasan (1967) proved that the constant-cycle property of the optimal plan extends to the case where demand is growing geometrically over time. D'Aspremont Gabszewicz and Vial (1972) embedded into the analysis the assumption of an exogenous depreciation rate, while Gabszewicz and Vial (1972) introduced the technological progress which can arise either at a deterministic date or at a random time in the future. In both cases, the constant-cycle property of the optimal policy still holds. A considerable amount of studies has been developed afterwards mainly in the operation research field, and variants of the capacity expansion model as it was formulated by Manne (1961) were used for several application to public services (see Luss 1982, Li and Tirupati 1992, and Nam and Logendran 1992 for a survey of the capacity expansion problem in OR and its main applications).

While in the context of developing countries, it was reasonable to assume that demand expansion was pushed exogenously by the population growth, today similar demand booms are observed for other reasons than an exogenous population growth. Indeed, they mainly derive from the globalisation of trade which makes industries progressively concerned with larger and larger geographical areas to be served.

Further, mainly in response to the failure of state-owned entities to meet the growing needs of customers, many public utilities have been progressively changed into private institutions, namely profit maximizing entities. As such, they are allowed to use according to a profit-maximization criterion specific instruments which were not available at the time when public utilities were required to satisfy the demand function, e.g. they can adopt a price policy in order to manipulate the demand function over time consistently with their installed capacity .

Accordingly, we propose to relax the assumption of exogenous demand increase and replace it by an assumption about the time expansion of a demand function specifying how instantaneous demand varies with price at each point of time. More precisely, we consider a monopoly producing a good whose intercept of its demand function grows linearly over time: at each price demand expands through time as the intercept of the demand function grows. Whatever the price that is quoted at each instant, the monopolist must invest in order to meet the resulting demand growth. Further, the monopolist can only invest in indivisible factory units. The problem we would like to study in this paper is the following. *What is the optimal price policy, and the ensuing optimal sequence of investment time points the monopolist should select through time ?* Of course, the answer to this question would be simple if the price selected by the monopolist would be kept constant over time. Indeed, in this case, the investment points would be fully determined by the length of the time-interval during which the capacity of the last installed factory is sufficient to face the demand increase at that price. Since the capacity of each factory is assumed to be constant, this in turn would imply that the investment points are equally spaced. Yet, the problem is no longer as simple if we consider at the same time the possibility opened to the monopolist to manipulate the price over time. In that case, he can either dampen demand increases by increasing the price trajectory or, on the contrary, benefit from full instantaneous monopoly profit by adapting simply the monopoly price to demand increase. Of course, these price manipulations will no longer keep necessarily the time interval between two investment points constant. For instance, if demand is dampened by a strongly increasing price trajectory after a given investment point of time, the next investment will have to be undertaken earlier than for a weaker price increase. Here, we characterise the optimal policy to be followed when the monopolist can manipulate the price trajectory and the resulting sequence of investment points under the constraint of meeting demand at each point of time. Our findings are as follows. The optimal price policy, through which the monopolist can possibly dampen or enhance the expansion of demand through time, leads to a sequence of time points when to install a new factory which are equally spaced. Furthermore, the optimal price pattern between two dates at which a new factory is installed falls into two categories: either the price pattern is such that total existing capacity which results from the number of installed factories is fully used at each instant between these dates; or the instantaneous monopoly price is used for some period within these dates and, thereafter, until the point of time when a new factory is installed, the price which dampens instantaneous demand at the level of installed factory is adopted. Finally, we show that the time between the points when new factories are installed is longer than the one which would have been selected by a planner facing demand levels corresponding to the instantaneous monopoly price, as in Manne (1961).

The model is described in section 2. In the same section, we fully characterise both the optimal price and investment policies. Then, we show how our model is related to other bulks of literature. We summarize our findings in the conclusion and propose some paths for further research.

3.2 Basic framework

We consider a monopolist facing a demand function $D(t, p)$ increasing over time defined as follows:

$$D(t, p) = At - p$$

where t denotes continuous time and p instantaneous price. At each price demand expands through time as the intercept of the demand function A grows proportionally to t.

The sequence of capacity increments x_{t_i} consists of a new factory whose size is constant and equal to $\bar{x}$. Then, at each instant of time t, the available capacity $X(t)$ is bounded by the existing number of installed factories $i\bar{x}_{t_i}$, where i denotes the number of factories which have been installed up to the time t. While the existing capacity may exceed the current demand level $D(t,p)$, no undercapacity is admitted, so that $D(t,p) \leq X(t)$. The investment cost for adding a new factory of size $\bar{x}$ is equal to:

$$f(x) = a\bar{x};$$

This cost structure is assumed to hold for ever (no technological progress). Also we assume no capital depreciation. Time is discounted at a constant interest rate r. The time horizon is unbounded. We assume that *(i) the sequence of capacity increments x_{t_i} consists of a constant investment size $\bar{x}$; (ii) the monopolist sets the price regime $p(t)$ for $t \in [t_i, t_{i+1}[$, and any period $[t_i, t_{i+1}[$ and, accordingly, selects the timing of investment $\{t_i\}, t_i = 0, 1, 2...$when to install the exogenously given capacity $\bar{x}$.* The dates selected by the monopolist in order to install new factories are called *regeneration points,* and the interval of time $[t_i, t_{i+1}[$ between two regeneration points is called *cycle.* In each cycle $[t_i, t_{i+1}[$ the existing capacity remains constant and is equal to $X(t_i) = \sum_{k:t_k \leq t_i} x_{t_k} = i\bar{x}$. In the usual plant size problem, an exogenous sequence of capacity increments would automatically determine the corresponding sequence of regeneration points. Yet, when the monopolist can manipulate the price through time, this correspondence ceases to operate since it depends on the price policy selected by her. A *price policy* is a function $p(t)$ which specifies the price announced by the monopolist at each instant

t. Given any price policy, we may associate to it a sequence of regeneration points $(t_1; t_2; ....t_i...) = \mathbf{t}(p(t))$. Formally the problem is to find $p(t)$ and, accordingly $\mathbf{t}(p(t)) = (t_1; t_2; ....t_i...)$, so that the objective function $V(\mathbf{t}, p(t))$

$$V(\mathbf{t}, p(t)) = \int_0^{\infty} p(t) D(t, p(t)) e^{-rt} dt - \sum_{t_i=0}^{\infty} a\bar{x}_{t_i} e^{-rt_i}$$

is maximized subject to the capacity constraint

$$D(t, p(t)) \leq X(t).$$

A policy is said to be *optimal* when it consists in an optimal price pattern through time and, as a consequence, an optimal sequence of dates when to install capacity in order to satisfy demand at each instant of time.

3.3 The Optimal Policy

In this section we start by considering the optimal price policy within a cycle. In order to find the optimal price pattern within a cycle we first remark that the cost function $\sum_{t_i=0}^{\infty} a\bar{x}_{t_i} e^{-rt_i}$, where the $t_i's$ correspond to that policy, does not depend on the choice of $p(t)$. Then, a sufficient condition for the optimality of $p(t)$ is that it maximizes the integrand $p(t) D(t; p(t))$ at any point t, given the capacity constraint.

Notice that within any cycle $[t_i, t_{i+1}[$, the objective function $V(\mathbf{t}, p(t))$ achieves its maximum for $p(t)$ given by

$$p(t) = \max(At/2, At - X_{t_i})$$

for $i = \max[n \mid t_n < t]$ the integer part of t. This follows from the maximization problem

$$\underset{p(t)}{max} = p(t)D\left(t; p\left(t\right)\right)$$
$$\text{s.t.} D\left(t; p\left(t\right)\right) \leq X_{t_i}, t_i = \left[t\right].$$

Whatever the selected sequence of regeneration points, during the cycle $[t_i, t_{i+1}[$, they may arise two price regimes, depending on the size of the factory $\bar{x}$ with respect to the demand. Assume that at some point of time t, where $t_i \leq t < t_{i+1}$, the capacity constraint is not binding, namely $D\left(t; p\left(t\right)\right) < X_{t_i}$. Then, at the optimal policy, $p(t)$ should be set equal to the maximizing price $p^M(t) = \frac{At}{2}$, as the demand does not need to be dampened. Yet, the demand expands over time while the current capacity remains fixed during the cycle $[t_i, t_{i+1}[$. When it happens that the capacity constraint turns out to be binding, namely $D\left(t; p\left(t\right)\right) = X_{t_i}$, then the firm has to choose the price $p^C(t) = At - X_{t_i}$ so as to contract the demand $D\left(t; p\left(t\right)\right)$ within the limits imposed by the existing factories. These two price patterns $p^M(t)$ and $p^C(t)$ are called, respectively, *monopoly price regime* and *constrained price regime.* Further, we denote by t_i^* the point of time when the monopoly price regime $p^M(t)$ becomes equal to the constrained price regime $p^C(t))$ and we label this point as the *switching point* t_i^*. It is easy to verify that $t_i^* = \frac{2X_{t_i}}{A}$. Then, three scenarios may arise. In the first scenario the switching point is exactly equal to t_{i+1}, so the optimal price pattern coincides with the monopoly price regime during the whole of the cycle (scenario A). In the second scenario , the switching point lies between the two investment dates of a cycle

$[t_i, t_{i+1}[$; then both these two regimes are used at the optimal price pattern, the first one between t_i and t_i^* and the second between t_i^* and t_{i+1} (scenario B). In the third scenario the switching point lies before t_i; then the monopolist is forced to use the constrained price regime during the whole of the cycle in order to meet the capacity constraint (scenario C). Figures 5 illustrate these 3 scenarios. We prove first that the switching point t_i^* can never be *exactly equal* to the regeneration point t_{i+1}, excluding thereby scenario A.

Proposition 11 *During any cycle $[t_i, t_{i+1}[$ the switching point t_i^* belongs either to the interior of the cycle t_i, or it is strictly smaller than t_i.*

Proof. Assume that, for an exogenously given size of the factory $\bar{x}$,$t^{'}_{i+1}$, $t^{'}_{i+1} \neq t_{i+1}$, would be the corresponding optimal decision point, and $t^{'}_{i+1} = t_i^*$. Then, the monopolist can quote the monopoly price regime during the whole of the cycle because the time length between t_i and $t^{'}_{i+1}$ is the one at which the factory installed at time t_i is equal to the demand level at time $t^{'}_{i+1}$, namely $X^{'}_{t_i} = D^M_{t^{'}_{i+1}}$. The present value of the discounted flow of revenues R during the cycle $[t_i, t_{i+1}[$ obtains as

$$R = \int_{t_i}^{t^{'}_{i+1}} \left(\frac{At}{2}\right)^2 e^{-rt} dt. \tag{24}$$

Assume now to postpone the investment of the new factory $\bar{x}$ for δ, where $t^{'}_{i+1} + \delta > t_i^*$. The firm gains the discounted cost $G = a\bar{x}_{t_i}(e^{-rt_{i+1}} - e^{-r(t_{i+1}+\delta)})$ saved by postponing the investment. Yet, the demand is not completely satisfied. This induces to switch from the monopoly price regime $p^M_{t_i}$ to the constrained price regime $p^C_{t_i}$. Accordingly, the present value of the discounted flow of revenues R_i during the cycle $[t_i, t^{'}_{i+1} + \delta[$ turns into

$$
\begin{aligned}
R' &= \int_{t_i}^{t'_{i+1}} \left(\frac{At}{2}\right)^2 e^{-rt} dt + \qquad (25)\\
&\int_{t'_{i+1}}^{t_{i+1}+\delta} \left[\left(\frac{A(t'_{i+1}+\delta)}{2} \right)^2 + \frac{A}{2}^2 (t'_{i+1}+\delta)(t-(t'_{i+1}+\delta)) \right] e^{-rt} dt,
\end{aligned}
$$

where the second integral denotes the revenue stemming from using the constrained price regime between t'_{i+1} and $t'_{i+1}+\delta$. Substracting (25) from (24) yields the loss L_{t_i} resulting from the switch between the two price regimes in the cycle $[t_i, t'_{i+1}+\delta[$:

$$
L = \int_{t'_{i+1}}^{(t'_{i+1}+\delta)} \left[\begin{array}{c} \left(\frac{At}{2}\right)^2 - \left(\frac{A(t'_{i+1}+\delta)}{2}\right)^2 + \\ -\frac{A}{2}^2 (t'_{i+1}+\delta)(t-(t'_{i+1}+\delta)) \end{array} \right] e^{-rt} dt
$$

This loss L is a function of the order of δ^2 over an interval of length δ, so its order is of magnitude δ^3, while the gain G is of first order in δ. Accordingly, for δ small enough the net loss is negative. Then if the investment in new factory would be postponed postpone, the corresponding discounted profit would increase, which is the desired contradiction. **Q.E.D.**

It follows from above that only two scenarios may be observed at the optimal price policy within a cycle: either *(i)* the level of demand at some $t \geq t_i$ can be satisfied by the existing equipment X_{t_i}, for any t_i, then the optimal price policy consists in quoting between the regeneration point t_i and the switching point t_i^* the monopoly price regime and, after the switching point up to the next regeneration point t_{i+1}, the constrained price regime; or *(ii)* the level of demand at t_i cannot be met by the existing equipment and then the constrained price regime is quoted within the whole of the cycle. Whether the monopoly price regime and the constrained price regime alternate within a cycle, or the constrained regime applies during the whole of the cycle, depends on the exogenous size of the factory $\bar{x}$ which is installed at the beginning of each cycle, and gain and costs resulting from investing at a certain date. If the the size of the new factory which is consented to be installed is large, the switching point t_i^* may result beyond the investment date t_i, and the optimal price policy involves the monopoly price regime between t_i and t_i^*, and the constrained regime from t_i^* up to t_{i+1} (Figure 6; case (i)) . On the contrary, if the size of the factory $\bar{x}$ is small, the optimal investment date t_i may result beyond the switching point t_i^*, and the optimal price policy results in quoting the constrained regime within the whole of the cycle (Figure 7; case (ii)).

Denote by x_0 the "fictitious" factory size such that $x_0 = D(p(0))$, and $t_i = t_i^*$. The following holds :

Lemma 12 *When $\bar{x}_{t_i} > x_0$, then $t_i^* > t_i$ and the optimal price policy consists in quoting both the price regimes within the cycle $[t_i, t_{i+1}[$, while for $\bar{x}_{t_i} < x_0$, $t_i^* < t_i$ and the constrained regime applies during the whole of the cycle.*

We prove now that both in the case when $\bar{x} > x_0$, and in the reverse one, the optimal investment time policy is unique and stationary. However, according to the above lemma, the corresponding optimal price policies differ in both cases.

Let us first remark the following, which will be used in the proof of the next propositions. Define the objective function $V(t)$ as

$$V(t) = \int_{t_i}^{t_{i+1}} \pi(t)e^{-rt}dt - a\bar{x}_{t_i}e^{-rt_i}$$

where $\int_{t_i}^{t_{i+1}} \pi(t)e^{-rt}dt = \int_{t_i}^{t^*} e^{-rt}\frac{(At)^2}{4}dt + \int_{t^*}^{t_{i+1}} e^{-rt}X_{t_i}(At - X_{t_i})dt$. Clearly the sequence of investment dates $t_i\,'s$ corresponding to the optimal price regimes must maximizes $V(t)$. This can be satisfied only when the cost $a\bar{x}_{t_i}$ and the revenue $\int_{t_i}^{t_{i+1}} \pi(t)e^{-rt}dt$ of installing a new factory exactly countervail.

We start now with the first case, and consider the case when $\bar{x} > x_0$.

Proposition 13 *In the case when the constant factory size $\bar{x} = \bar{x}_{t_i}$ exceeds the value of x_0 for which $t_i = t_i^*$, the optimal price policy involves both the monopoly and constrained price regimes ; furthermore, the resulting sequence of regeneration points is stationary.*

Proof. First, notice that the first part of the proposition follows immediately from the first part of the lemma. Consider a specific cycle $[t_i, t_{i+1}[$ and assume that the firm invests in a new factory at time $t_i + \delta$ instead of time t_i. Then, the cycle $[t_{i-1}, t_i[$ grows longer by δ, while the cycle $[t_i, t_{i+1}[$ turns into $[t_i + \delta, t_{i+1}[$. Thus, the change in the revenue function R due to prolonging the cycle $[t_{i-1}, t_i[$ and postponing for δ the starting date of the cycle $[t_i, t_{i+1}[$ is between t_i and $t_i + \delta$ and writes

$$\Delta R = \int_{t_i}^{t_i+\delta} \left(\frac{A^2 t^2}{4} - \frac{A^2}{4} [2t_{i-1}^* t - (t_{i-1}^*)^2] \right) e^{-rt} dt.$$

$\frac{A^2t^2}{4}$ is the equation of a parabola representing the revenue function if the factory x_{t_i} is installed at time t_i, while $\frac{A^2}{4}[2t_{i-1}^* t - (t_{i-1}^*)^2]$ is the equation of a straight line denoting the revenue function if the new factory is installed at time $t_i + \delta$ and the constrained price regime is quoted from the switching point t_{i-1}^* up to the end of the cycle.

The difference in the cost function C is

$$\Delta C = a\bar{x}_{t_i}(e^{-rt_i} - e^{-r(t_i+\delta)}) \sim \delta a\bar{x}_{t_i} e^{-rt_i}$$

up to terms quadratic in δ. Then the net change writes as the difference between the change in cost function and the change in the revenue function, namely:

$$(\Delta C - \Delta R) = e^{-rt_i}\delta[a\bar{x}_{t_i}r - \frac{A^2}{4}(t_i - t^*_{i-1})^2]$$

It is easy to see that given x_{t_i}, there is a unique t_i such that the net change cancels out. As this value of t_i only depends on the parameters of the model and the exogenous increment of capacity and all of them are constant over time, then the optimal policy is unique and stationary. **Q.E.D**.

We consider now the second case, namely when $\bar{x} \leq x_0$ and the switching point t^*_i is out of the cycle $[t_i, t_{i+1}[$, that is $t^*_i < t_i$, and we prove that even in this case the investment time policy which is induced by the price policy is unique and stationary.

Proposition 14 *In the case when the constant factory size $\bar{x} = \bar{x}_{t_i}$ is smaller than the value of x_0 for which $t_i = t^*_i$,the optimal price policy involves the constrained price regime only; furthermore, the resulting sequence of regeneration points is stationary.*

Proof. First, notice that the first part of the proposition follows immediately from the second part of the lemma. Consider again the specific cycle $[t_i, t_{i+1}[$. Suppose that firm installs the factory x_{t_i} at time $t_i + \delta$ instead of time t_i. The only change in the revenue function R due to postponing investment is between t_i and $t_i + \delta$. As $x_{t_i} \leq x_0$, then $t^*_i < t_i$, and the optimal price policy consists in quoting the constrained price regime within the whole of the cycle. Accordingly, the difference in the revenue function writes as follows:

$$\Delta R = \int_{t_i}^{t_i+\delta} \left(\frac{A^2}{4}[2t_{i-1}^* t - (t_{i-1}^*)^2 - \frac{A^2}{4}[2t_i^* t - (t_i^*)^2] \right) e^{-rt} dt$$

where the first term $\frac{A^2}{4}[2t_{i-1}^* t - (t_{i-1}^*)^2$ denotes the revenue function when the new factory is installed at time $t_i + \delta$ while the second term $\frac{A^2}{4}[2t_i^* t - (t_i^*)^2]$ denotes the revenue function when the new factory is installed at time t_i. Again, the difference in the cost function C is

$$\Delta C = a\bar{x}_{t_i}(e^{-rt_i} - e^{-r(t_i+\delta)}) \sim \delta r a \bar{x}_{t_i} e^{-rt_i}$$

up to terms quadratic in δ. Then the net change is

$$(\Delta C - \Delta R) = e^{-rt_i}\delta[a\bar{x}_{t_i} r - \frac{A^2}{4}(t_i^* - t_{i-1}^*)(2t_i - (t_i + t_{i-1}^*)]$$

So, given $\bar{x}_{t_i}$, and the parameters of the model there is a unique t_i for which the net change $(\Delta C - \Delta R)$ cancels out. **Q.E.D**.

Thus, regardless of the exogenous factory size, the sequence of regeneration points is stationary. Thus, the optimal price policy while affecting the trajectory of demand through time does not alter the stationarity property of the investment pattern.

Let us now show how at the optimal price policy the regeneration point t_i depends on the parameters of the model. First, notice that the regeneration point t_i is set in such a way that the gain which derives from saving the discounted costs exactly counterbalances the loss due to the gap between the existing factories and the demand level corresponding to the monopoly regime. Then, we can write the following:

$$a\bar{x}_{t_i}r = \frac{A^2}{4}(t_i - t^*_{i-1})^2 \tag{26}$$

Further, we know from the lemma that when $\bar{x}_{t_i} < x_0$ the switching point t^*_i is out of the cycle $[t_i, t_{i+1}[$; when the reverse holds, then the switching point t^*_i is in the interior of the cycle $[t_i, t_{i+1}[$; finally when $\bar{x}_{t_i} = x_0$ the switching point t^*_i and the regeneration point t_i coincides. Accordingly, when the factory size $\bar{x}_{t_i}$ is at the fictitious level x_0, the condition (26) can be re-expressed as:

$$ar\frac{A}{2}(t_i - t^*_{i-1}) = \frac{A^2}{4}(t_i - t^*_{i-1})^2$$

where $\bar{x}_{t_i} = X_{t_i} - X_{t_{i-1}}$, $X_{t_{i-1}} = \frac{A}{2}t^*_{i-1}$ and $X_{t_i} = \frac{A}{2}t^*_i = \frac{A}{2}t_i$. Then, easy computations yield:

$$t_i = \frac{2ar}{A} + t^*_{i-1}.$$

and the following holds:

Lemma 15 *At the optimal price policy, the value of each regeneration point is higher the higher the cost of investing (a and r) and the lower the growth parameter of demand A.*

Furthermore, it follows from the above findings that the optimal price policy entails a sequence of cycles which are longer than the ones which would have been selected by a planner facing demand levels growing exogenously at a rate corresponding to the instantaneous monopoly price. If the monopolist would quote at each instant of time the monopoly price, it would generate a linear demand trajectory with slope $\frac{A}{2}$ in the demand-time space. We can find out what is the optimal solution of the planning problem corresponding to this expansion of demand, viewed as in Manne (1961). To this solution, there corresponds an optimal sequence of equally-spaced regeneration points. Starting with this sequence of cycles, we can examine whether the optimal price policy which is adopted by the monopolist leads to cycles which is larger or smaller than the ones resulting from the planning solution provided by Manne.

Proposition 16 *The optimal price policy entails a sequence of cycles which are longer than the ones which would have been selected by a planner facing demand levels corresponding to the instantaneous monopoly price.*

Proof. First, notice that at the optimal price regime, the instantaneous monopoly price is never quoted for the whole of the cycle. Then, the demand function can be either increasing from the beginning of the cycle up to the switching point, and then dampened up to the end of the cycle, when the optimal price policy entails alternating both the monopoly and the constrained regimes, or dampened for the whole of the cycle when the constrained regime is only quoted. Similarly, we have seen that during the cycle resulting from the optimal regime, the factory size cannot exceed or coincide with the demand level $D^M\left(t_{i+1}, p^M(t_{i+1})\right)$. Then, at the optimal cycle the factory size must be strictly smaller than $D^M\left(t_{i+1}, p^M(t_{i+1})\right)$ while, at the Manne solution, it is exactly equal to this magnitude. Accordingly, given the size of the factory, at the optimal price policy, the cycle during which demand levels are satisfied is longer then the one for which demand levels resulting from the instantaneous monopoly price are served. **Q.E.D.**

3.4 Relation with the investment literature

This paper is also related to other bulks of literature. First, it is concerned with the application of microeconomic models of discontinuous adjustments (retail inventories, expansion of capacity, cash balances, prices, durable goods etc.) to the aggregate economic behavior. These microeconomics models were formulated in the fifties for retail inventories, and then applied for a variety of different economic problems, such as labor demand, durables goods and s.o. Recently, starting from them, a new approach has been developed with the aim of investigating aggregate implications of lumpy adjustments of equipment at a plant level (Blinder 1981, Caplin 1985, Caplin and Spulber 1987, Caballero and Engel 1991, *inter alia*). In particular, on the basis of Caplin and Spulber (1987), Caballero and Engel (1991) have developed a methodology which allows to explain why the economic behaviour of the aggregate investment displays a smooth pattern at the aggregate level, while individual production units adjusting discretely and by large amount according to a (S,s) policy. Although this approach has been conceived within the framework of (S,s) policy, it can still be applied, with minor changes, to our model and thus used for evaluating macroeconomic implications of investment policy in term of capacity expansion as it is defined at a plant level.

Second, the stationarity property we find also holds in models concerned with the optimal timing of replacement investments. These replacement models assume that the technological progress is embodied in new equipment and accordingly examine when it is optimal replacing the oldest capital with the most recent vintages. The endogenous scrapping time, namely the age of the oldest machines, which is determined by the replacement timing selected by firms, is proved to be constant over time, under several different conditions (Malcomson 1975, Van Hilten 1991, Boucekkine, Germain and Licandro 1997). Furthermore, when this optimal scrapping rule is used to examine phenomena which precisely result from replacing old machines, such as job creation and job destruction, a cyclicity property still holds even in models concerned with these phenomena (Boucekkine, del Rio and Licandro 1999). Although our specification, as it is, does not focus on the replacement problem, our modeling fits into the literature on replacement decision. With slightly changes, we could assume that a factory is required to replace its equipment which deteriorates with new machines rather then to expand its capacity to meet a rising demand, and then evaluating at the optimal price policy the ensuing timing of replacement.

3.5 Conclusion

In this paper, we have derived the optimal investment timing and pricing policies when the size of the investment is fixed. It represents a natural extension of the literature devoted to the capacity expansion problem which has flourished during the last decades. It brings a new element into the picture: the fact that the firm can manipulate its price trajectory through time in order to accelerate, or to dampen, demand expansion. We have proved that the introduction of this new element does not alter the basic property of the optimal investment trajectory, as derived in the capacity expansion literature, namely the constant-cycle property. Furthermore we have shown that the optimal cycle is longer than the cycle corresponding to demand expansion which results from the monopoly price regime. There is at least a further direction on which to build upon our model, namely considering the replacement problem which can be faced by a factory when its equipment deteriorates. Although this topic have been largely examined in a specific field of growth theory, the so called vintage literature, nevertheless it could be

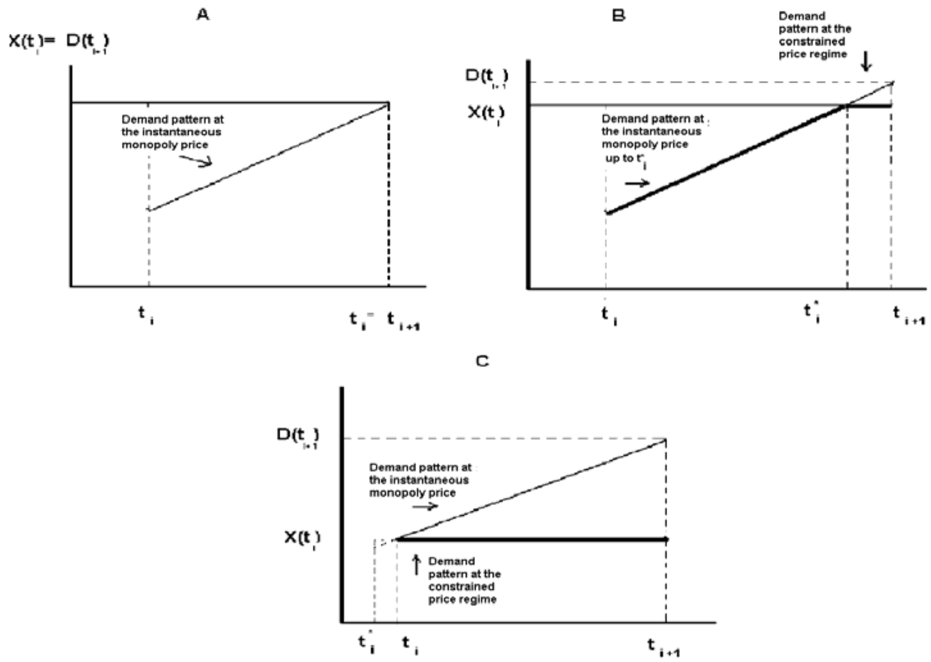

Figure 5: Demand function in 3 different price regime scenarios.

interesting to nest this problem in our model as specific case in order to evaluate it from a microeconomic view point.

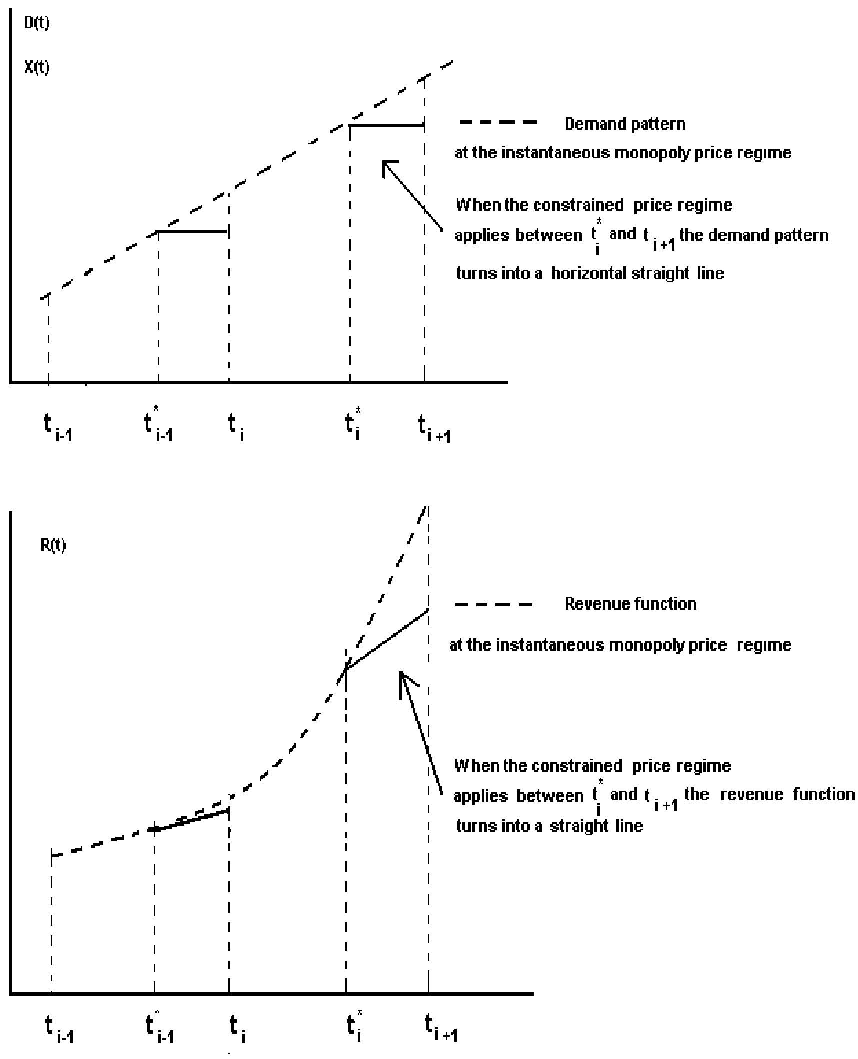

Figure 6: Demand function and revenue function when the optimal price regime involves alternating both the monopoly and the constrained regimes.

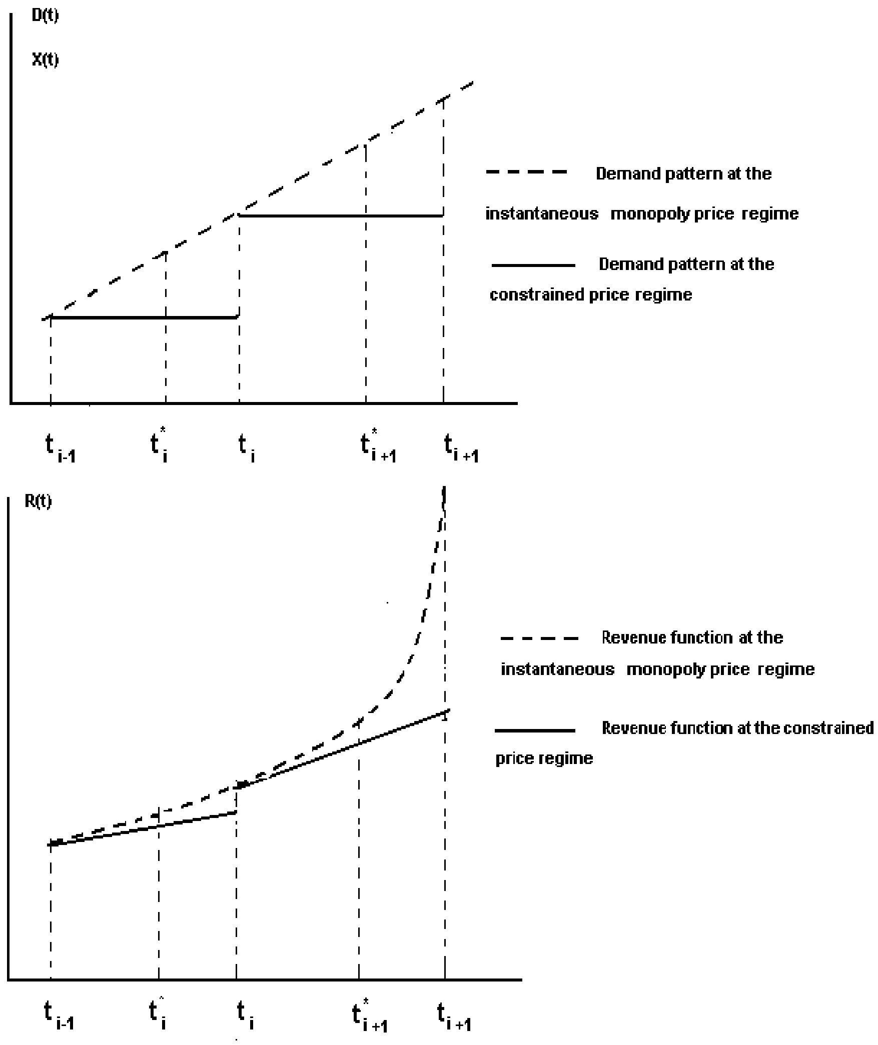

Figure 7: Demand function and revenue function when the optimal price regime only consists of the constrained regime.

References

Abel A.B. and J.C. Eberly (1996). "Optimal Investment with Costly Reversibility", *Review Of Economic Studies* 63, 581-593.

Acemoglu D. (1999). "Changes in Unemployment and Wage Inequality: an Alternative Theory and Some Evidence", *American Economic Review* 89, 1259-78

Acemoglu D.(2002) "Technical Change, Inequality and The Labor Market", *Journal of Economic Literature* XL, 7-12.

Agarwal R. and M. Gort (2000). "First Mover Advantage and the Speed of Competitive Entry, 1887-1986", Manuscript, SUNY at Buffalo.

Aghion P. and P. Howitt (1992). "A Model of Growth Through Creative Destruction", *Econometrica* 60, 323-351.

Aghion P. and P. Howitt (1998). Endogenous Growth Theory. Cambridge: MIT Press.

Aghion, P., Howitt, P., Harris, C., and J. Vickers (2001), "Competition, Imitation, and Growth with Step-by-Step Innovation", *Review of Economic Studies* 28, 467-92.

Allen S. (2001) "Technology and the Wage Structure", *Journal of Labor Economics* 19, 440-483.

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

Boucekkine R. Del Rio F.and O. Licandro (1999). "Endogenous vs Exogenously Driven Fluctuations in Vintage Capital Models", *Journal of Economic Theory* 88*,* 161-187*..*

Boucekkine R. Germain M. and O. Licandro (1997). "Replacement Echoes in the Vintage Capital Models", *Journal of Economic Theory* 74*,* 333-348

Caballero R. J. and Engel E (1991). "Dynamic (S, s) Economies", *Econometrica* 59, 1659-1686.

Caballero R.J. and E.M.R. Engel (1999)."Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach", *Econometrica* 67, 783-826.

Chaouch B.A. and J.A. Buzacott (1994). The effects of lead time on plant timing and size, *Production and Operations Management*, 3(1), 38-54.

Caplin A. S. and Spulber D.F. (1987). "Menu Costs and the Neutrality of Money", *The Quarterly Journal of Economics* 102, 703-726.

Chenery H. (1952). "Overcapacity and the Acceleration Principle", *Econometrica*

Choi J.P. (1989). "An Analysis of cooperative R&D", W.P., Harvard University.

Cozzi G. (1999). "R&D Cooperation and Growth", *Journal of Economic Theory 86*, 17-49.

D'Aspremont C., Gabszewicz J.J. and J.P.Vial (1972). "Capital depreciation, growing demand and economies of scale", *CORE DP* No. 6918.

D'Aspremont C. and Jacquemin A. (1988). "Cooperative and Noncooperative R & D in Duopoly with Spillovers", *American Economic Review* 78, 1133-1137.

Dinopoulos E. and P. Segerstrom. (1998). "A Schumpeterian Model of Protection and Relative Wages," *American Economic Review* 89, 450-472.

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

Gabszewicz J.J. and J.P.Vial (1972) Mathematical Methods in Investment and Finance, G.P. Szego and K. Shell eds., *Optimal Capacity Expansion Under Growing Demand and Technological Progress*, 159-189.

Green J.R. and Scotchmer S. (1995) "On the Division of Profit in Sequential Innovation, *RAND Journal of Economics* 26, 20-33.

Greenwood J., Hercowitiz Z. and Krussel P. (1997). "Long-Run Implications of Investment-Specific Technological Change", *American Economic Review* 87, 342-62.

Greenwood J. and Yorucoglu (1997). "1974", Carnegie-Rochester Conference Ser. Public Pol. 46, 49-95.

Grossman G.M. and Helpman E. (1991). "Quality Ladders in the Theory of Growth", *Review of Economic Studies* 58, 43-61.

Howitt P. (1999). "Steady Endogenous Growth with Population and R & D Inputs Growing", *Journal of Political Economy* 107, 715-730.

Jones, C. (1995), "R&D Based Models of Economic Growth", *Journal of Political Economy* 103: 759-84.

Jones C. (1999). "Growth: With or Without Scale Effects?", *American Economic Review* 89, 139-144.

Jovanovic B. and G. MacDonald (1994). “Competitive Diffusion”, *Journal of Political Economy* 102, 24-52.

Jovanovic B. and R. Rob (1989). “The Growth and Diffusion of Knowledge”, *Review of Economic Studies* 56, 569-582.

Jovanovic B. and P Rousseau (2001) “Mergers and Technological Change: 1885-1998” *WP, N01/16*, Vanderbilt University..

Kamien M. I., Muller E. and Zang I. (1992) “Research Joint Ventures and R&D Cartels”, *American Economic Review* 82, 1293-1306.

Kortum S. (1997). “Research Patenting and Technological Change”, *Econometrica* 65, 1389-1419.

Kremer M. and Maskin E. (1999). “Segregation by Skill and the Rise in Inequality”. Harvard Mimeo.

Li S. and D. Tirupati (1992). "Technology Selection and Capacity Expansion for Manufacturing Systems", chapter 10 in Intelligent Design and Manufacturing, A. Kusiak (ed.), J. Wiley, New York.

Link A., D. Paton and D. Siegel, (2002), "An Analysis of Policy Initiatives to Promote Strategic Research Partnerships", *Research Policy*, 31, 1445-1457.

Luss H. (1982) " Operations Research and Capacity Expansion Problems: A Survey", *Operations Research* 41, 5, 907-947.

Malcomson M. (1975). "Replacement and the Rental Value of Capital Equipment subject to Obsolescence", *Journal of Economic Theory* 10, 24-41.

Manne A. S. (1961). "Capacity Expansion and Probabilistic Growth", *Econometrica.*

Manne A. S. (1967). (Ed.), Investments for Capacity Expansion, London.

Nam S. and R. Logendran (1992). "Aggregate Production Planning: a Survey of Models and Methodologies", *European Journal of Operation Research* 61, 255-272.

NicKell S. (1977) "Uncertainty and lags in the investment decision of firms", *Review of Economic Studies* 44, 249-263.

Norhaus W.D (1969) "An Economic Theory of Technological Change", *American Economic Review* 59, *18-28.*

O'Donoghue T. (1998). "A Patentability Requirement for Sequential Innovation", *RAND Journal of Economics* 29, 654-679.

O'Donoghue T., Scotchmer s. J. Thisse (1998). ""Patent Breadth, Patent Life, and the Pace of Technological Improvement", *Journal of Economics and Management Strategy* 7, 1-32

Peretto P. (1998). "Technological Change and Population Growth", *Journal of Economic Growth 3,* 283-311

Ryan S. M. (2002). "Capacity Expansion for Random Exponential Demand Growth with Lead Times", *W.P.* Dep. of Industrial Manufacturing System Engineering, Iowa State University.

Shapiro C. (2002). "Competition Policy and Innovation", *Working Paper* 2002/11, DSTI.

Scotchmer S (1996). “Protecting Early Innovators: Should Second-Generation Products be Patentable?”, *RAND Journal of Economics* 27, 322-331.

Scotchmer S (1999). “On the Optimality of the Patent Renewal System”, *RAND Journal of Economics* 30, 181-196.

Schumpeter J. A. (1942). Capitalism, Socialism and Democracy. New York: Harper & Row.

Segerstrom P. (1998). “Endogenous Growth without Scale Effects”, *American Economic Review* 88, 1290-1310.

Srinivasan T. N (1967). ”Geometric rate of growth of demand”, in: Investments for capacity expansion, op.cit., chapter 9.

Tarola, O. (2002). “Essays in Schumpeterian Growth Theory”, PhD Dissertation, University of Rome ”La Sapienza”.

Van Hilten (1991) ”The Optimal Lifetime of Capital Equipment”, *Journal of Economic Theory* 55, 449-454.

Young A. (1998). “Growth Without Scale Effects”, *Journal of Political Economy* 106, 41-63.