Empirical patterns of ¯rm growth and R&D investment: a quality ladder model interpretation

by

Tor Jakob Klette¤
(University of Oslo and Statistics Norway)

and

Zvi Griliches
(Harvard University and NBER)

August 25, 1998

Abstract

We present a model of endogenous ¯rm growth with R&D investment and stochastic innovation as the engines of growth. The model for ¯rm growth is a partial equilibrium model drawing on the quality ladder models in the macro growth literature, but also on the literature on patent races and the discrete choice models of product di®erentiation. We examine to what extent the assumptions and the empirical content of our model are consistent with the ¯ndings that have emerged from empirical studies of growth, productivity, R&D and patenting at the ¯rm level. The analysis shows that the model ¯ts well with empirical patterns such as (i) a skewed size distribution of ¯rms with persistent di®erences in ¯rm sizes, (ii) ¯rm growth independent of ¯rm size, as stated in the so-called Gibrat's law, and (iii) R&D investment proportional to sales, as well as a number of other empirical patterns.

1 Few theories on firm growth and firm heterogeneity

Empirical research on firm and plant level data has revealed a large amount of heterogeneity within narrowly defined industries. This heterogeneity is striking in a number of dimensions such as size (sales, employment), firm growth rates, rates of job creation and job destruction, and also in variables such as capital intensity and R&D intensity[1]. Much of the heterogeneity, e.g. in terms of sales and R&D investment, is quite persistent over a number of years. Recently, researchers have addressed the question of how we can reconcile this persistent heterogeneity with theories based on optimizing agents (see the survey by Sutton, 1997). This paper presents a model of endogenous firm growth where R&D investment and stochastic innovation are the engines of growth. Throughout the analysis, we examine to what extent the model is consistent with the empirical evidence on firm growth and the findings from the microeconometric research on R&D, innovation and patents.

The model presented below can be made consistent, for specific parameter values, with at least three widely studied empirical regularities of R&D investment, firm sales and firm growth: (i) R&D intensities are independent of sales[2]. This result is derived from our model which treats R&D investment as a non-rival input in production. (ii) Firm growth is, to a first approximation, independent of size. This relationship is often referred to as Gibrat's law[3]. (iii) The size distribution of firms is highly skewed with persistent differences in firm sizes. This is true both for sales and other variables such as R&D investment. This third regularity is closely related to the second regularity, as has been emphasized by Simon and his co-workers (see Ijiri and Simon, 1977). Our contribution is to show that these three regularities can all be related and derived from a fully specified model of endogenous firm growth, based on optimizing agents. We examine whether the empirical content of the assumptions required to derive the model make sense in view of what has been learned from microeconometric research on innovation, patents and R&D. We also consider a number of additional empirical implications that we derive from the model.

Our objective is to develop a model that can serve as a framework for firm level analysis of R&D investment, innovation, patenting and firm performance in terms of profitability and growth. In the present paper, we discuss the empirical implications of our model in relation to results from microeconometric studies presented elsewhere. The model presented below is an

---

[1]Some recent contributions to this empirical research are e.g. Dunne et al. (1988), Mairesse and Griliches (1990), and Davis and Haltiwanger (1992). The large literature on the skewed distribution of firm sizes is surveyed in Sutton (1997).

[2]See Cohen and Klepper (1996) for a recent review of this issue.

[3]See Sutton (1997) for a comprehensive survey.

"extreme hypothesis" in Simon's (1968) terminology, and we do not, of course, expect the model to fit all relevant empirical facts[4]. The model can hopefully serve as a benchmark for empirical analysis, and we discuss some extensions of the framework in the final section.

The specification of the model is inspired by the macro-models of endogenous growth, in particular the quality ladder model developed by Grossman and Helpman (1991a,b) and Aghion and Howitt (1992). Their version of the quality ladder model implies that each new innovation is introduced by a new firm. Barro and Sala-i-Martin (1995, ch.7) and Thompson (1996) point out that it is hard to reconcile this property with the observed pattern with persistent dominance of established firms, at least on an annual time scale. We have therefore developed an alternative model where incumbent firms persistently innovate and grow or decline over a number of years. The relationship between our model and the previous literature on the quality ladder model is briefly discussed in appendix A, where we also list a number of other studies related to our analysis. Ericson and Pakes' (1995) study has a focus similar to ours in that they develop a theoretical model of R&D investment and firm growth through innovation consistent with a number of empirical observations. Another related study is Cohen and Klepper (1996) which provides an interpretation of the empirical patterns of R&D investment similar to our model, but they do not address the issue of firm growth. See appendix A for further remarks on these and other studies.

Compared to the early work by Herbert Simon and others on firm growth and Gibrat's law[5], our framework gives some value added in that the random walk process for the firms' demand is not imposed on the model as an a priori assumption, but emerges as an endogenous result from quality changes as a result of R&D activities which are treated in the model as forward looking investment behavior.

The next section spells out the model and discusses the validity of the assumptions of the model in view of existing empirical studies. Additional empirical implications of the model are derived and examined in section 3. Section 4 elaborates on the specification of the demand side of our model, the nature of price competition and the optimal price setting, which are treated rather briefly in section 2. Section 5 provides conclusions and discussion of future research tasks.

---

[4]This issue reminds us of a quotation from Picasso: "Art is the lie that help us see the truth". We accept that our model - as any model - is "a lie", but we also hope that it can help us to reveal true insights about firm growth and performance.

[5]See e.g. Ijiri and Simon (1977) and the survey by Sutton (1997).

2 Our version of the quality ladder model

2.1 The quality ladder model, macro growth and ¯rm growth

As pointed out above, our analysis is a variation of the quality ladder model, with elements from the patent race literature concerned with the persistence of monopoly. We have replaced the speci¯cation of the demand side in the quality ladder models of Grossman-Helpman (1991a, 1991b) and Barro-Sala-i-Martin (1995, ch.7), with a speci¯cation based on the discrete choice theory for di®erentiated products. Our alternative speci¯cation of the demand side allows the model to account for both the growth and the decline of ¯rms.

2.2 Alternative speci¯cation of the demand side

Consider a ¯rm producing a product \i" of quality k_i and with price P_i.

Assumption 1: The ¯rm's output is determined by the demand function

$$Q_i = M \frac{e^{v(k_i; P_i)}}{\sum_{j \in J} e^{v(k_j; P_j)}} \tag{1}$$

with $v(k_i; P_i) = {}^\circ k_i \text{ ¡ } {}^\circledR \ln P_i{}^6$. $^\circ$ and $^\circledR$ are parameters and M is the size of the market. J refers to all the competitors in the market (including foreign competitors).

This demand system has been extensively discussed by others; see e.g. Anderson, dePalma and Thisse (1992) and Berry (1994) and references cited in those papers. Through this demand system, ¯rm growth (or decline) in our model is driven by cumulative improvements in product quality of the ¯rm's own product (k_i) as well as improvements in the product quality of competing products. These cumulative improvements in product quality is generated by R&D investments. R&D investments are driven by expected pro¯ts, and a ¯rm's pro¯t increases with the size of the ¯rm and therefore by its relative product quality (cet.par.).

As just stated, the model focuses on quality k_i as a determinant of demand and it is product innovations, i.e. the changes in k_i, generated by the ¯rm's R&D investment that determines ¯rm growth. It is useful to specify the demand function as $Q_i = Q(k_i; {}^a)$, i.e. as a function of the product quality k_i, while lumping the other variables together in the variable ^{a} [7]. These other variables (^{a}) include the ¯rm's product price and the prices and qualities of the competing

---

[6] The speci¯cation of v can easily be extended so that $v(k_i; P_i; !_i) = {}^\circ k_i \text{ ¡ } {}^\circledR \ln P_i + !_i$, where $!_i$ captures unobservable demand shifters, apart from quality, as e.g. in Berry (1994). This would not a®ect the analysis presented below.

[7] In the analysis below, ^{a} will contain the denominator in the expression on the left hand side of ([?]), which include k_i. We will, however, assume that the number of competitors, i.e. J, is su±ciently large so that the ¯rm will neglect its own impact on this denominator.

products (cf. equation 1), and the ¯rm considers ^{a} as determined independently of the level of its own k_i. In section 4, we discuss the properties of the demand system (1) in more detail and its implication for optimal price setting behavior. In particular, we discuss to what extent it is consistent to assume that the price of a product is independent of its quality.

2.3   Dynamic optimization and ¯rm growth

Let us now focus on one particular di®erentiated product, and the process of quality upgrading through R&D of this product. Each step in this upgrading involves a three stage game. In the ¯rst stage, the incumbent determines R&D investment in its product line. In stage two, there is free entry by outside competitors into the product line. When the new upgraded version of the product is developed, the game enters the third stage, with price competition between all the di®erentiated product lines (¯rms) in the industry, conditional on the prevailing product qualities. In this setup, we should emphasize the assumption:

Assumption 2: The incumbent decides its level of R&D investment before potential entrants.

As clari¯ed in the exchange between Gilbert-Newbery and Reinganum in the American Economic Review, this assumption about the order of moves has empirical content; see Gilbert and Newbery (1984). It is required to favor the persistent dominance of the incumbent. We noticed above that there is a large empirical literature that has documented the persistency of leading ¯rms[8], at least with a focus on observation covering some years rather than decades. The process of \creative destruction", where one dominant ¯rm is replaced by another, is emphasized both by Reinganum (1985) in her model of industry evolution, and in the quality ladder models by Grossman-Helpman (1991a, 1991b) and Aghion-Howitt (1992). Such a process of creative destruction might be more relevant for the longer run, i.e. in terms of decades rather than years. Hall, Griliches and Hausman (1986) have documented the persistency of di®erences in R&D e®ort between ¯rms in the same industry[9]. We will return to this question below.

We will now present a formal analysis of optimal R&D investment that leads us to a model of ¯rm growth. Consider one particular ¯rm producing a given product with a certain quality level. In our model, quality improvements for this product will take place through R&D e®ort. The ¯rm with its product of quality k earns pro¯t $¦(k)$ per unit of time. We will suppress the ¯rm subscript and the other variables entering the pro¯t function to avoid notational clutter; these other variables include prices and quality variables for the competitors. Assume that the

---

[8]Cf. Sutton (1997). See also the discussion in Gruber (1992) and Klepper (1996).

[9]One should notice that the persistency of R&D e®ort documented in Hall, Griliches and Hausman (1986) is somewhat exaggerated, since their analysis is based on a continuing, and hence heavily selected, panel.

product is replaced by a new product after a period $T(k)$, where $T(k)$ is a stochastic variable. It follows that the value of getting the new product, conditional on $T(k)$, is given by

$$V(k;T) = \mu(k)(1 - e^{-rT(k)})/r \tag{2}$$

The arrival time of the next product $(k+1)$ is uncertain. Denote the probability of the arrival of the next product per unit of time by $a(k)$, which is some function of R&D investment and the quality of the existing product, i.e. $a(k) = h(R;k)$. We expect $h_1 > 0$, while we will argue below that h_2 is negative. The cumulative distribution function for $T(k)$ is exponential: $P(T(k) < \tau) = G(\tau) = 1 - \exp[-a(k)\tau]$. It follows that the density $g(\tau) = G'(\tau) = a(k)\exp[-a(k)\tau]$. The expected value of the product k is therefore given by

$$
\begin{aligned}
E_T \, V(k;T) &= \mu(k)a(k)/r \int_0^\infty (1 - e^{-r\tau}) e^{-a(k)\tau} d\tau \\
&= \frac{\mu(k)}{r + a(k)} \tag{3}
\end{aligned}
$$

The incumbent firm is potentially replaced by an outside competitor producing a superior product denoted by $k + 1$ [10]. The expected reward for participating in the race for product $k + 1$ per unit of time is

$$\frac{a(k)\,\mu(k+1)}{r + a(k+1)}, \tag{4}$$

where $a(k+1)$ is the probability per unit of time for the arrival of product $k + 2$. (4) shows that cet.par. a larger probability of innovation (i.e. higher a's) has two opposing effects on the expected rewards to invest in innovation in this model. A larger probability of innovation increases the probability to innovate in the first round which increases the expected returns, but it also increases the probability that the innovation will be replaced with a subsequent innovation, thereby reducing the expected return. With positive discounting, (4) shows that the first effect dominates.

In order to preserve its dominance, the incumbent will carry out just sufficient research to deter outsiders to enter the R&D race. We assume that there is free entry into research and that there is a sunk cost associated with entering the race[11]. The zero profit condition requires

---

[10] Notice that just a small set up cost for new entrants within the existing product line will ensure that no entrants will enter the market with a product k, since the Bertrand competition will drive out all profits. Our model does not explicitly examine entry into new product lines as this type of entry will affect all firms in the industry in a symmetric way, through shifts in the variable we have denoted a. We want to focus primarily on firm heterogeneity with any given industry. See further remarks below.

[11] As pointed out by Gilbert and Newbery (1984) in a similar analysis, a free entry equilibrium is only well defined in the presence of sunk cost in R&D. Sunk costs in R&D is also an essential part of Cohen and Klepper's (1996) analysis of R&D investment.

that the net present value of participating in the race is equal to the sunk cost (F), i.e.

$$\int_0^\infty e^{-[r+\delta(k)]t}[\delta(k)EV(k+1) - R]\,dt$$

$$= \frac{\delta(k)\,EV(k+1) - R}{r + \delta(k)} = F(R) \tag{5}$$

The notation $F(R)$ reflects that sunk cost is allowed to depend on the level of the R&D activity. To be able to solve out for the optimal level of R&D, we must make an additional assumption about the functional form for the innovation-function $h(R;k)$:

Assumption 3: Probability of project success and quality improvement per unit of time, is given by

$$\delta = R\,A(k): \tag{6}$$

with $A'(k) < 0$.

Assuming constant returns with respect to R in the innovation function as in (6) is analytically very convenient, but empirically questionable[12]. The functional form for the innovation function $h(R;k)$ is, of course, related to the issue of \diminishing returns in R&D" that has been the subject of much research on the basis of patent statistics; see Griliches (1990) for a survey. In a study of the relationship between patents and R&D in a panel of U.S. firms, Hall, Griliches and Hausman (1986) estimated the elasticity to between 0.3 and 0.6 in the longitudinal dimension, suggesting rather sharply diminishing returns. However, the appearance of diminishing returns in the longitudinal dimension could be an artifact due to the incompleteness of the underlying data rather than a reflection of the characteristics of the underlying innovation process, as discussed in Griliches (1990).

Even if the estimated elasticity of patents with respect to R&D expenditures had been precisely estimated in the patent studies, it is not clear how that should be translated to the form of the innovation function $h(R;k)$. That is, what is the relationship between a patent and making a step on the quality ladder; does a patent increase demand with a certain percentage or is the percentage increase in demand from a patent dependent on the stock of patents and the firm's size. This is a another functional form question that remains open[13]. We will return to the discussion of the functional form for the innovation function $h(R;k)$ below.

---

[12] The linear specification in R is also problematic since δ, being a probability, should be restricted to the interval between 0 and 1. Hence, this linear specification should be considered an approximation to the (empirically) relevant segment of e.g. an S-shaped δ-function restricted to the 0-1 interval.

[13] The relationship between sales and patents might be so noisy that it is hard to discriminate between alternative functional forms on the basis of the available data. The standard assumption in the literature is to assume a log-log relationship between patent counts and sales, rather than a linear-log relationship; see e.g. Griliches, Hall and Pakes (1991). Exploring the functional form of the relationship between patents and sales is an issue that we hope to examine in future research.

The sunk cost in R&D, $F(R)$, is assumed to increase with the magnitude of the research e®ort; larger R&D activity requires e.g. more laboratory facilities and other setup costs. For simplicity, we will assume that the sunk cost increases linearly with the size of the R&D e®ort, i.e. $F(R) = {}^\circ R$: With this speci¯cation and (6), it follows from (5) that the incumbent's R&D e®ort is determined by

$$R = \frac{EV(k+1)}{^\circ} \mathbf{i} \; \frac{1}{\acute{A}(k)} \left(\frac{1}{^\circ} + r\right) \tag{7}$$

Pro¯ts are assumed to increase linearly with output: $\frac{1}{4}(k) = \frac{1}{4} Q(k)$, which implies that the pro¯ts per unit of output is not a®ected by the ¯rm's product quality. In section 4, we examine to what extent this assumption is consistent with optimal price setting behavior in our model of demand, and we also discuss how the model can be extended to allow for innovation a®ecting prices and pro¯ts per unit of output. Inserting the expression $\frac{1}{4}(k) = \frac{1}{4} Q(k)$ and (3) into (7), we get

$$R(k) = \frac{\frac{1}{4}Q(k+1)}{^\circ[r + {}_{\scriptstyle ,}(k+1)]} \mathbf{i} \; \frac{1}{\acute{A}(k)} \left(\frac{1}{^\circ} + r\right) \tag{8}$$

2.4 Firm growth according to Gibrat's law

We are now ready to show under what conditions the model will generate ¯rm growth rates independent of ¯rm size, according to Gibrat's law. In order for Gibrat's law to hold in this model, we must have some diminishing returns in the innovation process:

Assumption 4: Making product improvements gets more and more di±cult, such that

$$\acute{A}(k) = \acute{A}_0 \, e^{\mathbf{i}\,\acute{}\,k}: \tag{9}$$

with $\acute{} > 0$:

The literature surveyed in Cohen and Klepper (1996) suggests that $\acute{A}$ is a declining function of k (see Cohen and Klepper's Stylized fact 4). $\acute{A}^0 < 0$ re°ects \¯shing out" e®ects where it gets harder and harder to make new innovations. This is another aspect of the \diminishing returns to R&D" issue that we also discussed above.

Some empirical evidence might be interpreted as support for Assumption 4. The study by Bound et al. (1984) found that the number of patents per dollar of R&D is signi¯cantly lower for ¯rms with larger R&D budgets[14]. Similarly, Acs and Audretsch (1991) found a negative

---

[14]Griliches (1990) discusses alternative interpretations of this ¯nding by Bound et al., and he argues that the pattern might be an artifact due to a sample selectivity problem for small ¯rms.

relationship between innovations per R&D dollar and the level of R&D investment[15]. In the time series dimension at the aggregate level, several researchers have emphasized a related pattern; i.e. the increasing ratio of R&D per patent, see Caballero and Jaffe (1993), Kortum (1993) and Griliches (1994). One possible interpretation of these findings is that they reflect the property stated in Assumption 4, as discussed also in Cohen and Klepper (1996).

On the other hand, both Assumptions 3 and 4 on the functional form of the innovation function must be considered as tentative rather than as empirical facts at this stage[16]. At this stage, Assumptions 3 and 4 are driven as much (or more) out of analytical convenience as from empirical support. We believe that further research is needed on the functional form issues, both research concerned with ways to alter or relax our assumptions in the theoretical analysis, and empirical research that tries to tie down the functional form of the innovation function[17].

Let us, with a slight abuse of notation, write $Q(k) = a \, e^{\sigma k}$, where a captures the other factors apart from the firm's own quality, that affect the firm's demand (as in section 2.2). This simplified expression for $Q(k)$ is only valid if each firm has a negligible impact on the denominator in the demand function (1). It follows by inserting this expression for $Q(k)$ and (9) into (8), that the R&D intensity can be written

$$\frac{R}{Q} = \frac{e^{\sigma \frac{1}{4}}}{(r+\gamma)^{\sigma}} \cdot \left[\frac{1}{\bar{A}_0^a} \frac{1}{\sigma} + r \right] e^{(\gamma - \sigma)k} \tag{10}$$

In the rest of this paper we will focus on the particular case when $\gamma = \sigma$, which is the case where the positive size effect on the firm's optimal R&D-intensity is just offset by the increased difficulty of improving the product as its quality is increased. In this special case we have the interesting property that γ is independent of k as we show formally in Appendix B. As just mentioned, there are two offsetting forces that makes γ independent of k. On the one hand, innovation is assumed to get harder and harder when k increases, as spelled out in Assumption 4, while on the other hand, the incentives to invest in R&D grow with the level of the firm's expected profits. Since future expected profits is proportional to the firm's current size in our model, we have that the incentives to invest in R&D increases with firm size, as shown in (8). With the functional form assumptions made above and $\gamma = \sigma$, the two effects just offset each other, with the result that γ turns out independent of k[18]. It follows that the whole right hand

---

[15]Acs and Audretsch's result is based on a cross sectional analysis of a comprehensive data set on innovations collected by the Small Business Administration.

[16]Another question that we have not discussed above, is whether it is correct to associate the longitudinal results as related to the R part in the $h(R; k)$-function, while the cross sectional results is associated with the k-part.

[17]See also Griliches (1994) for a broader argument reaching a similar conclusion.

[18]This result will hold for more general functional forms than considered above; the essential condition is that the product $Q(k)\acute{A}(k)$ is independent of k.

side of (11) is independent of k. When $\acute{} = °$, (10) can be written

$$\frac{R}{Q} = \frac{e^° \frac{1}{4}}{(r + \,)^°} \; i \; \frac{1}{\acute{A}_0 \, a} \left[\frac{1}{°} + r \right]$$ (11)

Equation (11) has a number of implications that we will examine in the next section. Let us ¯rst summarize the model: The ¯rms in the industry compete with products that are both vertically and horizontally di®erentiated and each ¯rm upgrades its product through a step-wise stochastic R&D process in order to preempt its product line from being overtaken by a potential competitor. As discussed above, the model is speci¯ed so that a ¯rm with a higher quality product will gain larger pro¯ts from an additional innovative step, but to make an additional innovative step is also more costly for a ¯rm with a higher quality product. The ¯rst feature suggests that higher quality ¯rms will invest more in R&D, while the other part pulls in the opposite direction.

Notice that a ¯rm operating in this industry will change its market share because of its own innovation and due to innovations in competing ¯rms. Since the competing ¯rms persistently improve their products, the ¯rm has to innovate to preserve its share of industry output.

3 Main propositions derived from the model and the empirical evidence

3.1 Results related to ¯rms within the same industry

Let us for the moment neglect di®erences in production costs and assume that di®erences in ¯rm size is largely due to di®erences in product quality[19]. Then, since the right hand side of (11) does not involve k, we have that

Proposition 1 R&D increases proportionally with ¯rm size so that R&D intensity is independent of size.

The empirical literature that relates to this issues is vast as is clear e.g. from the survey by Cohen and Klepper (1996)[20]. Bound et al. (1984), in their study based on the Compustat ¯le, concluded after checking a number of econometric issues, that R&D increases proportionally to size. There were deviations from this pattern among very large and very small ¯rms, which tended to be more R&D intensive than the rest. However, as they point out, very small ¯rms on the Compustat ¯les are likely to be more innovative and do more R&D than the average small

---

[19]We should warn the reader that there is a distinction between two dimensions of quality that we will clearify in section 4.3 below.

[20]See also the surveys by Cohen and Levin (1989) and Cohen (1995).

¯rm in US manufacturing[21]. Cohen et al. (1987) con¯rmed the main conclusion in Bound et al. at the ¯rm level, and also at the line of business level for the sample of R&D performing business units. They report, however, some positive relationship between the R&D intensity and the size of the business unit for the sample of all business units. In their survey, Cohen and Levin (1989) emphasize that the size e®ects in the R&D intensities, even in cases where they are found to be statistically signi¯cant, are \minute both in terms of the variance explained and the magnitude of the coe±cients". Similarly, Stylized fact 3 in Cohen and Klepper (1996) states that \in most industries it has not been possible to reject the null hypothesis that R&D varies proportionally with size across the entire ¯rm size distribution". We conclude that Proposition 1, and hence the model presented in this paper, is a reasonable ¯rst approximation to at least one widely studied pattern of R&D investment.

To the extent di®erences in productions costs and prices are signi¯cant determinants of di®erences in ¯rm sizes, Proposition 1 needs to be restated. α, capturing the demand e®ect of di®erences in prices, appears on the right hand side of (11), and one can show that this factor will tend to create a positive relationship between the R&D intensity and ¯rm size. If di®erences in prices can be captured by so-called \¯xed e®ects", the ¯rst part of Proposition 1 still holds, i.e. R&D increases proportionally with ¯rm size in the longitudinal dimension.

Equation (11) also implies that:

Proposition 2 The R&D intensity is positively related to the pro¯t per unit of output.

Two studies that empirically examine the relationship between R&D intensity and the pro¯t margin are Geroski, Machin and van Reenen (1993) and Brouwer and Kleinknecht (1994)[22]. Both studies ¯nd a positive and statistically signi¯cant relationship. Brouwer and Kleinknecht's study is a pure cross sectional analysis, while the study by Geroski et al. uses a set of panel data. The study by Geroski et al. emphasizes a causal relationship that runs from changes in R&D to changes in the pro¯t margin, but they also point out that the longitudinal e®ects are small compared to persistent cross sectional di®erences in the pro¯t margins. In the analysis presented above, the causal relationship runs from pro¯ts per unit of output to R&D; products with higher expected pro¯ts per unit of output will cet.par. attract more R&D investment. Further research comparing the cross sectional and the time series relationship between pro¯t margins and R&D might be able to clarify the causal relationship, but we recognize that it is notoriously hard to

---

[21]Griliches (1990) elaborates on this argument.

[22]The pro¯t margin, de¯ned as $(P - C^0)/P$ where C^0 is marginal costs, is not exactly the same as pro¯ts per unit of output, but at least with constant marginal costs, movements in the pro¯t margin will have a positive and direct relationship to shifts in pro¯ts per unit of output.

identify the causal structure by regression analysis. In section 4, we will show how our model can be extended to accommodate cases where an innovation increases pro¯ts per unit of output, i.e. with higher pro¯ts per unit of output for recent innovations and where the pro¯ts per unit of output gradually decreases as the innovations get older. This extension allows also for a causal relationship from R&D to higher pro¯t margins, as identi¯ed by Geroski et al. If the extended model is correct, we should not be able to identify a one-way causal relationship between R&D and the pro¯t margins, since the causality runs both ways.

Above we argued that ¸ is independent of k; see Appendix B for the formal argument. Since k determines the size of the ¯rm and ¸ is the Poisson parameter in the stochastic process for k, we have established the following proposition[23]:

Proposition 3 The model implies ¯rm growth independent of ¯rm size, in accordance with Gibrat's law.

There is a long line of research on the empirical relationship between ¯rm size and ¯rm growth; see the survey by Sutton (1997). Two relevant studies are Evans (1987) and Hall (1987). Evans (1987) concludes that departures from Gibrat's law might be signi¯cant for small ¯rms and long time periods, but one \might not go too wrong by maintaining Gibrat's law" for \short run changes in the growth and size distribution of the largest ¯rms". Hall (1987) has carried out a careful analysis considering sampling bias and measurement errors, and she concludes that \Gibrat's law is weakly rejected for the smaller ¯rms ... and accepted for the larger ¯rms". The evidence suggests that Gibrat's law is not universally true, and it would not make sense to take that pattern for granted. However, we consider it desirable that our model can be made consistent with this benchmark case.

From Propositions 1 and 3:

Proposition 4 R&D follows a random walk (Gibrat's law in R&D).

In a study of US manufacturing ¯rms (over 8 years), Hall, Griliches and Hausman (1986) con-cluded that \R&D investment [is] essentially a random walk with an error variance which is small relative to the total variance of R&D expenditures between ¯rms". The persistency of cross-sectional di®erences in R&D has also been documented in Klette and Johansen (1998).

It is well known that growth according to Gibrat's law will generate a skewed size distribution of ¯rm size:

---

[23]This result assumes that price di®erences are unrelated to di®erences in the level of k. Section 4 examines how this can be consistent with optimal price setting behavior.

Proposition 5 The model will generate a highly skewed distribution of firm size.

To our knowledge, there has been little systematic analysis of the skewness of the size distribution of high-tech industries (or other industries for that matter) since the book by Ijiri and Simon (1977). Figure 1 displays the size distribution for business units in the 8 most R&D intensive (4-digit) industries in Norwegian manufacturing. The width of the graphs in Figure 1 corresponds to the number of business units in the industry. The graphs are based on log of sales for the business units, and the boxes show the location of the quartiles in the distribution for each industry, together with the median shown as the horizontal line inside each box. The whiskers indicate the spread of the distribution in the lower and upper quartiles. Consider e.g. the industry ISIC 3832[24] which has the largest box in figure 1, indicating that it is the largest industry. This box suggests that the size distribution in this industry is even more skewed than the log-normal distribution, since the median is closer to the lower quartile than to the upper quartile, and also since the upper whisker is longer than the lower whisker. Considering the other boxes in Figure 1, we find that most of the other industries also have a size distribution at least as skewed as the log-normal distribution.

Since the intensity parameter (λ) for the arrival of new innovations (i.e. increases in k) is independent of the level of k, the process will generate a Poisson distribution for k with a parameter which increases in proportion with the \age" of the process[25]. Since we consider the firms (in, say, an industry) to be different realizations of such processes, we have that:

Proposition 6 The model will generate a more and more widely spread distribution of firm sizes as the average firm age increases.

The increasing spread generated by Gibrat's law has been the focus of much research; see Steindl (1968) for a survey of the early attempts to create models for the firm size distribution, based on random walk growth but with a stable (non-increasing) spread of the distribution, and McCloughan (1995) for a recent study and further references. Entry and exit of firms can be one stabilizing force, as was early demonstrated by Simon and his coworkers; see Ijiri and Simon (1977). Notice that our model predicts an increasing spread of the firm size distribution only if the average firm age increases. The issue of average firm age is, of course, largely determined by the pattern of entry and exit, which is left open in the model presented above[26].

---

[24] ISIC 3832 is the manufacturing of communication equipment.

[25] The parameter in the Poisson distribution is proportional to (λt) in the case with a stationary intensity parameter (λ), where t is the age of the process. More generally, with a non-stationary intensity parameter, we have that the spread (and the mean) in the Poisson distribution is $\int_0^t \lambda(s)\, ds$:

[26] Pakes (1994) considers structural analysis of firm growth incorporating entry and exit decisions.

3.2 Di®erences in R&D intensities between industries

While entry (more precisely, preemption) within each product line is an endogenous part of our model, we do not explicitly model the entry and exit of competing product lines (cf. J in equation 1). It is consequently problematic to use our model, in particular (11), to explain di®erences in R&D between industries, as e.g. entry conditions typically di®er across industries. More generally, the problem is that since industries di®er, say, in terms of innovative opportunities, we would also expect ^{a} on the right hand side of (11) to di®er.

If we, for argument's sake, abstract from this problem and treat ^{a} as an independent parameter, we can use (11) to explain di®erences in R&D intensities between industries. The right hand side of (11) includes parameters related to innovative opportunities ($Á_0$), the consumer willingness to pay for higher quality products ($°$) and other factors determining demand (^{a}). Equation (11) shows that, according to our model:

Proposition 7 Industries with higher innovative opportunities, larger willingness to pay for higher quality products and larger demand will have higher R&D intensities.

These results are intuitive and there is a large body of research supporting Proposition 7; see the survey in Cohen (1995, section 4)[27]. Our model suggests how observables corresponding to these parameters could be incorporated in a structural analysis of R&D investment and ¯rm growth across industries. We will add a few comments on this issue in the ¯nal section.

4 The demand side and price setting

4.1 Stronger ¯rm interaction: Growth and decline of ¯rms

The standard quality ladder model captures little in terms of competition between available products. This is our primary motivation for introducing an alternative speci¯cation of the demand side. In our speci¯cation, based on discrete choice theory for the demand of di®erentiated products, it is relative product quality that determines the ¯rms' market shares, as well as relative price. A ¯rm facing demand speci¯ed as in (1) has to upgrade the quality of its product at the pace of its competitors in order to preserve its relative size, as we noticed above.

In addition to competition from a new, innovating ¯rm that potentially can overtake its product line, a ¯rm is facing price competition from two di®erent margins: (i) Competition from other di®erentiated products as captured by the denominator in (1), and (ii) competition from the lower quality versions of the same di®erentiated product. The latter case, where the

---

[27]See also Schmookler (1966) and Pakes and Schankerman (1984).

price determination of a new product is constrained by competition from the lower quality versions of the same differentiated product, corresponds to a non-drastic innovation in Arrow's (1962) terminology, adopted in the patent race literature[28]. Otherwise the innovation is termed drastic.

4.2 Optimal price setting and the profit margin

Non-drastic innovations

Let us denote the marginal costs for firm i by C_i^0. The lower quality variety of product i (denoted $k_i - 1$) can be produced by competing firms, and we assume for simplicity that they have the same marginal costs as firm i. If the monopoly price for firm i exceeds $\ln C_i^0 + \gamma/\beta$, the new (high-quality) product will be outcompeted by the old variety[29]. This is the case of a non-drastic innovation. In this case, as shown by Arrow (1962), the optimal price is the so-called \limit price", determined such that the lower quality product is just competed out of the market, i.e.

$$\ln P_i = \ln C_i^0 + \gamma/\beta \qquad (12)$$

minus a little epsilon. Clearly, this price is determined independently of k_i, and hence the profit per unit of output is independent of k_i and therefore the size of the firm as we assumed above.

Drastic innovations

In the case of drastic innovations, a firm with a new innovation can charge the monopoly price as it is not constrained in its price setting by the presence of lower quality varieties of the same product, i.e. its monopoly price is below $\ln C_i^0 + \gamma/\beta$. The price in the case with a drastic innovation is determined from the first order condition

$$\frac{\partial}{\partial P_i}\left[\left(P_i - C_i^0\right) Q_i\right] = \qquad (13)$$

$$Q_i + (P_i - C_i^0)\frac{\partial Q_i}{\partial P_i} = 0: \qquad (14)$$

With Bertrand competition and demand as specified in (1), we have that $\partial Q_i/\partial P_i = -\beta Q_i(1 - s_i)/P_i$, and hence

$$P_i = \left(1 - \frac{1}{\beta(1 - s_i)}\right)^{-1} C_i^0 \qquad (15)$$

---

[28]The terms drastic and non-drastic innovations are usually applied to situations with process innovations, but they are equally applicable to product innovations with the demand system specified as in our model.

[29]This follows from the demand system (cf. eq. 1), as consumers will choose the low quality variety at price C_i^0 if $v(k_i; P_i) < v(k_i - 1; C_i^0)$:

where s_i is the firm's share of industry output, i.e. $s_i \equiv Q_i/M = Q_i / \sum_{j \in J} Q_j$ [30]. In terms of profits per unit of output, (15) can be rewritten

$$\mu_i = P_i - C_i^0 = \frac{C_i^0}{\beta(1 - s_i) - 1} \qquad (16)$$

Figure 2 shows how profit per unit of output, as expressed on the right hand side of (16), varies with firm size for different values of β. Figure 2 reveals that for moderate market shares and not too small price elasticities (cf. β), the profit per unit of output increases slowly with the firm's market share, i.e. is determined largely independent of k_i.

To sum up, the optimal price setting rule in our model implies that profit per unit of output is independent of market shares in cases with non-drastic innovations, and as an approximation also in many cases with drastic innovations. The empirical evidence on the relationship between profit margins, market shares and firm size has been surveyed by Schmalensee (1989); see in particular his Stylized facts 4.11 and 4.12 and the subsequent discussion. Griliches and Cockburn (1994) found that increased competition and drastic changes in market shares due to entry of new generic drugs in the pharmaceutical industry they considered did not seriously affect the price charged by the incumbent firm. To the extent that the shift in market share did not affect the marginal production costs, this finding is consistent with the price setting in our model. Metrick and Zeckhauser (1996) provide some empirical evidence suggesting that the profit margins are independent of product quality in the markets for mutual funds and in the automobile industry. There is, however, also some indirect counter evidence e.g. in the study by Geroski, Machin and van Reenen (1994), as discussed in relationship to Proposition 2.

Profit margins and the product life cycle

Empirical studies, e.g. Mansfeld, Schwartz and Wagner (1981), suggest that the profit margin might be high for a new high-quality product variety, but that imitation eliminates innovative rents within a few years. Similarly, Berndt, Griliches and Rosett (1993) found that the rate of price change for a product is related to the age of the product[31]. We will now show how the model can be extended to incorporate a negative relationship between the profit margin and the age of the innovation.

Available evidence suggests, as mentioned, that the profit margin declines gradually after a new product is introduced. This pattern can be captured in our model in the case of non-drastic innovations, if we assume that the marginal costs of potential producers of the old variety is

---

[30] That $M = \sum_{j \in J} Q_j$ can be seen by summing both sides of (1) across all firms.
[31] See also Kuznets (1930).

gradually declining, e.g. as the knowledge of how to produce the lower quality variety di®uses to low cost producers. A declining pro¯t margin can also be explained in this way for innovations that are drastic when the new variety is introduced. If the production costs for the old, lower quality variety is (rapidly) declining, the pricing of products of drastic innovations might also turn out to be constrained by the competition of lower quality varieties some time after the new, high quality variety was introduced.

Formally we can assume that the pro¯ts decline with the age of the product variety according to $\pi_0 e^{-\frac{1}{2}\tau}$, where τ is the time elapsed since the most recent version of the product was introduced[32]. With this slight respeci¯cation we ¯nd that (2) should be rewritten $V(k) = \pi_0(k)(1 - e^{-(r+\frac{1}{2})T(k)})/(r+\frac{1}{2})$: It follows that (3) then will be replaced by

$$E_T V(k;T) = \frac{\pi_0(k)}{r + \frac{1}{2} + \delta(k)};$$ (17)

where $\pi_0(k)$ is the pro¯ts of the ¯rm when the current product variety was ¯rst introduced. Extending the model in this way does not require any substantial changes to the analysis, and it allows for a causal relationship running from R&D to innovation to pro¯t margins (cf. our discussion of Proposition 2), with the ratio between price and marginal costs being a declining function of the age of the most recent product variety. However, pro¯ts per unit of output remain independent of the level of the product quality in this extended model, which is essential to preserve the simple structure of our model of endogenous R&D and innovation as spelled out in section 2.

4.3   R&D as a non-rival input and product quality

So far we have assumed that productions costs are independent of the product quality. That is, our model assumes that once a new product improvement is developed and introduced, no additional resources are needed to produce this improved product as compared to the older version. Romer (1990) has emphasized this aspect of R&D, by labeling R&D a non-rival input of production. A number of cases can surely be listed suggesting that R&D is a non-rival input to production, but Adams and Ja®e (1996) and Klette (1996) have presented evidence suggesting that R&D is not a completely non-rival input at the ¯rm level. This can be a question of the level of aggregation and R&D might still be a non-rival input at the business or product line level.

---

[32] It would be preferable to consider the case where pro¯t per unit of output rather than total pro¯ts, declines exponentially over time, i.e. $\mu_0 e^{-\frac{1}{2}\tau}$, but the expressions are more complicated and will not be presented here.

Our framework can be considered to involve two di®erent dimensions of quality, one dimension which is explicit in our model, captured by the index k_i, and a second dimension which is hidden in our model. The second dimension of quality can be altered by spending more resources per unit of production { this quality aspect is a rival output dimension; it is related to the notion of a rival good in Romer's (1990) terminology. More horsepower and leather seats in cars are examples of quality di®erences in \rival" dimensions.

The ¯rst dimension { emphasized in this paper { is the non-rival aspect of quality; as soon as you know how to produce goods of a higher quality in the k_i-dimension, you do not have to spend more resources per unit of output for producing the higher quality product. A faster processor for computers can serve as an example of a non-rival improvement in quality.

We have implicitly buried the rival dimension in the measurement of output quantities and prices, and in the marginal costs. It is not surprising that the costs related to the rival part of product development and research, increases with sales. The more interesting part of our model is the non-rival aspect of product development and research, which distinguishes R&D from other factors of production.

5   Conclusions and remaining issues

We have presented a fully speci¯ed model of endogenous ¯rm growth, where R&D and innovations are the engines of growth. The model is tightly speci¯ed and is based on a number of ad hoc assumptions. However, these assumptions are not arbitrary; they were introduced in order to rationalize a number of empirical regularities that have been established from empirical research on ¯rm growth and innovation. We have also examined the empirical content of the assumptions and argued that they are good or acceptable approximations to the ¯ndings in much available empirical research.

On the basis of our analysis, we have argued that our model is promising as a benchmark model to understand:

1. Why the size distribution of ¯rms is highly skewed, with persistent heterogeneity.

2. How Gibrat's law can be reconciled with optimizing behavior.

3. Why R&D intensity is largely independent of size, even in cases where R&D is largely a non-rival input.

Our model is able to address these three issues within a single, integrated framework.

We must admit that the correspondence between the empirical literature and the empirical content of our model is not perfect. For instance, our model in some respects ignores imitation and other sources of innovation and quality changes[33]. Empirically, it is widely observed that even in high tech industries there is a large fraction of ¯rms reporting no R&D activity. These ¯rms presumably survive by imitation[34]. Another incomplete aspect of the model is that it leaves considerable heterogeneity to be rationalized. For instance, we observe empirically large and persistent di®erences in R&D intensities between ¯rms in most industries also at a dis-aggregated industry level. The sources of di®erences in R&D intensities have been subject of much research[35], and our model provides some handles to capture such heterogeneity, such as di®erences in pro¯t margins and di®erences in innovative opportunities. Identifying and mod-eling these di®erent sources for heterogeneity is a crucial step towards structural estimation of the parameters in the model.

---

[33]Notice, however, that the threat of entry which plays an important role in the model, implicitly assumes that imitation of existing products and know-how is easy since a new competitor potentially can make a step on the quality ladder from the same level as the incumbent ¯rm.

[34]See e.g. Nelson (1988) for a discussion of this point.

[35]See Pakes and Schankerman (1984) and Cohen (1995).

6 References

Acs, Z.J. and D.B. Audretsch (1991): R&D, ¯rm size, and innovative activity. In Acs, Z.J. and D.B. Audretsch (eds.): Innovation and Technological Change: An international Comparison. New York: Harvester Wheatsheaf.

Adams,J. and A. Ja®e (1996): Bounding the e®ects of R&D: An investigation using matched establishment-¯rm data, RAND Journal of Economics 27, 700-22.

Aghion, P. and P. Howitt (1992): A model of growth through creative destruction, Econometrica 60, 323-51.

Andersen, S.P., A. de Palma, J.-F. Thisse (1992): Discrete Choice Theory of Product Di®erentiation, Cambridge (U.S.): MIT Press.

Arrow, K.J. (1962): Economic welfare and the allocation of resources to inventions, in R.R. Nelson (ed.): The rate and direction of inventive activity, Princeton: Princeton University Press.

Barro, R. and X. Sala-i-Martin (1995): Economic Growth, New York: McGraw-Hill.

Berndt, E.R., Z. Griliches and J. Rosett (1993): Auditing the producer price index: Micro evidence from prescription pharmaceutical preparations. Journal of Business & Economic Statistics, 11, 251-64.

Berry, S.T. (1994): Estimating discrete choice models of product di®erentiation, RAND Journal of Economics 25, 242-62.

Bound, J., C. Cummins, Z. Griliches, B. Hall, A. Ja®e (1984): Who does R&D and who patents? in Z. Griliches (ed.): R&D, Patents and Productivity, Chicago: University of Chicago Press

Brouwer, N.M. and A. Kleinknecht (1994): Monopoly pro¯ts from innovation and wage di®erentials. A micro-econometric investigation. Mimeo, Free University of Amsterdam.

Caballero, R.J. and A. Ja®e (1993): How high are the giants shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In O.J. Blanchard and S. Fischer (eds.): NBER Macroeconomic Annual 1993. Cambridge (U.S.): MIT Press.

Cohen, W.M. (1995): Empirical studies of innovative activity. In P. Stoneman (ed.): Handbook of Innovation and Technological Change. Oxford: Blackwell.

Cohen, W.M. and S. Klepper (1996): A reprise of size and R&D, Economic Journal 106, 925-51.

Cohen, W.M. and R. Levin (1989): Empirical studies of innovation and market structure. Ch. 18 in R. Schmalensee and R. Willig (eds.): Handbook of Industrial Organization, Vol. 2. Amsterdam: North Holland.

Cohen, W.M., R. Levin and D. Mowery (1987): Firm size and R&D intensity: A re-examination. Journal of Industrial Economics, 35, 543-65.

Dasgupta, P. (1986): The theory of technological competition, in J.E. Stiglitz and G.F. Mathewson (eds.): New developments in the analysis of market structure. Cambridge (U.S.): MIT Press.

Davies, S. and J. Haltiwanger (1992): Gross job creation, gross job destruction, and employment reallocation, Quarterly Journal of Economics 107, 819-64.

Dunne, T., M. Roberts and L. Samuelson (1988): Dynamic patterns of ¯rm entry, exit and growth, RAND Journal of Economics 19, 495-515.

Evans, D.S. (1987): The relationship between ¯rm growth, size, and age: Estimates for 100 manufacturing industries, Journal of Industrial Economics, 35, 567-81.

Ericson, R. and A. Pakes (1995): Markov perfect industry dynamics: A framework for empirical analysis, Review of Economic Studies 62, 53-82.

Geroski, P.A., S. Machin, J. Van Reenen (1993): Innovation and ¯rm pro¯tability, RAND Journal of Economics 24, 198-211.

Gilbert, R.J. and D.M.G. Newbery (1982): Preemptive patenting and the persistence of monopoly, American Economic Review 72, 514-26.

Gilbert, R.J. and D.M.G. Newbery (1984): Uncertain innovation and the persistence of monopoly: Comment, American Economic Review 74, 238-42.

Griliches, Z. (1994): Patent statistics as economic indicators: A survey. Journal of Economic Literature 28, 1661-707.

Griliches, Z. (1994): Productivity, R&D, and the data constraint, American Economic Review 84, 1-23.

Griliches, Z. and I. Cockburn (1994): Generics and new goods in pharmaceutical price indexes, American Economic Review 84, 1213-32.

Griliches, Z., B.H. Hall and A. Pakes (1991): R&D, patents, and market value revisited: Is there a second (technological opportunity) factor?, Economics of Innovation and New Technology 1, 183-201.

Grossman, G.M. and E. Helpman (1991a): Innovation and Growth in the Global Economy, Cambridge (U.S.): MIT Press

Grossman, G.M. and E. Helpman (1991b): Quality ladders in the theory of economic growth, Review of Economic Studies, 58, 43-61.

Gruber, H. (1992): Persistence of leadership in product innovation. Journal of Industrial Economics, 40, 359-75.

Hall, B.H. (1987): The relationship between ¯rm size and ¯rm growth in the US manufacturing sector, Journal of Industrial Economics, Vol.. 35, 583-606

Hall, B.H., Z. Griliches and J.A. Hausman (1986): Patents and R&D: Is there a lag? International Economic Review 27, 265-83.

Ijiri, Y. and H. Simon (1977): Skewed size distributions and the sizes of business ¯rms. Amsterdam: North Holland

Jovanovic, B.(1982): Selection and the evolution of industry, Econometrica 50, 649-70.

Kihlstrom, R.E. and J.J. La®ont (1979): A general equilibrium theory of ¯rm formation based on risk aversion, Journal of Political Economy 87, 19-48.

Klepper, S. (1996): Entry, exit, growth, and innovation over the product cycle. American Economic Review, 86, 562-83.

Klette, T.J. (1996): R&D, scope economies and plant performance, RAND Journal of Economics 27, 502-22.

Klette, T.J. and F. Johansen (1998): Accumulation of R&D capital and dynamic ¯rm performance: a not-so-¯xed e®ect model. Annales d'Economie et de Statistique. (Forthcoming)

Kortum, S. (1993): Equilibrium R&D and the decline in the patent-R&D ratio: U.S. evidence. American Economic Review (Papers and Proceedings) 83, 450-57.

Kuznets, S. (1930): Secular movements in production and prices, Boston: Houghton Mi² in Co.

Lippman, S.A. R.P. Rummelt (1982) Uncertain imitability: An analysis of inter¯rm di®erences under competition, Bell Journal of Economics 13, 418-38.

Lucas, R.E. (1978): On the size distribution of ¯rms, Bell Journal of Economics 9, 508-23.

J. Mairesse and Z. Griliches(1990): Heterogeneity in panel data: Are there stable production functions?, in P. Champsaur et al. (eds.): Essays in honor of Edmond Malinvaud, Vol.. 3: Empirical economics. Cambridge (US): MIT Press.

Mans¯eld, E. M. Schwartz and S. Wagner (1981): Imitation costs and patents: An empirical study. Economic Journal, 91, 907-18.

McCloughan, P. (1995): Simulation of concentration development from modi¯ed Gibrat growth-entry-exit processes. Journal of Industrial Economics, 43, 405-33.

Metrick, A. and R. Zeckhauser (1996): Price versus quantity: Market clearing mechanisms when sellers di®er in quality. NBER Working Paper 5728.

Nelson, R.R. (1988): Modeling the connections in the cross section between technical progress and R&D intensity. RAND Journal of Economics 19, 478-85.

Pakes, A. (1994): Estimation of dynamic structural models. Problems and prospects II: Mixed continuous-discrete models of market interactions. In C. Sims and J.J. La®ont (eds.): Proceedings of the 6th World Congress of the Econometric Society. Cambridge (U.K.): Cambridge University Press.

Pakes, A. and R. Ericson (1990): Empirical implications of alternative models of ¯rm dynamics. NBER Working Paper 2893.

Pakes, A. and P. McGuire (1994): Computing Markov-perfect Nash equilibria: numerical implications of a dynamic di®erentiated product model, RAND Journal of Economics 25, 555-89.

Pakes, A. and M. Schankerman (1984): An exploration into the determinants of research intensity, in Z. Griliches (ed.): R&D, Patents and Productivity, Chicago: Chicago University Press .

Reinganum, J.F. (1985): Innovation and industry evolution, Quarterly Journal of Economics 100, 81-98.

Romer, P.M. (1990): Endogenous technological change, Journal of Political Economy 98, S71-S102.

Schmalensee, R. (1989): Inter-industry studies of structure and performance. Ch. 16 in R. Schmalensee and R. Willig (eds.): Handbook of Industrial Organization, Vol. 2. Amsterdam: North Holland.

Schmookler, J. (1966): Invention and economic growth. Cambridge (US): Harvard University Press.

Simon, H.A. (1968): On judging the plausibility of theories. In B. van Rootselaar and J.F. Staal (eds.): Logic, Methodology and Philosophy of Sciences, Vol.. III. Reprinted in Ijiri and Simon (1977).

Steindl, J. (1968): Size distributions in economics. In D.L. Sills (ed.): International Encyclopedia of Social Sciences. London: MacMillan.

Sutton, J. (1997): Gibrat's Legacy. Journal of Economic Literature 35. 40-59.

Thompson, P. (1996): Technological opportunity and the growth of knowledge: A Schumpeterian approach to measurement. Journal of Evolutionary Economics, 6, 77-97.

7 Appendix A: Previous literature

The specification of our model is inspired by the macro-models of endogenous growth, in particular the quality ladder model developed by Grossman and Helpman (1991a, 1991b) and Aghion and Howitt (1992). Most versions of the quality ladder model have the property that each innovation is introduced by a new firm, while the producers of the old product varieties are driven out of the market. It is difficult to reconcile this property with the persistency of large firms that we tend to see in empirical data, at least at an annual frequency. Thompson (1996) has made the same criticism of the quality ladder growth model. He develops a complete model of R&D investment and firm growth, somewhat different from the model we present. On the basis of his model, he presents an empirical analysis of the relationship between R&D and the stock market value of the firm.

Our model is similar to the model presented in Barro and Sala-i-Martin (1995, ch.7) that emphasizes the persistent dominance of the established firms, rather than the continuous replacement of the leading firm. There is, however, a serious error in the model by Barro and Sala-i-Martin (1995, ch. 7.4), and this is why we have developed our own model. Specifically, the second order condition associated with the first order condition in equation (7.39) is not satisfied. That is to say, the level of R&D investment that Barro and Sala-i-Martin consider for the incumbent is not the optimal level of R&D investment.

Our (partial equilibrium) version of the quality ladder model is closely related to patent race models discussing the persistence of monopoly; see in particular Gilbert and Newbery (1982, 1984). Reinganum's (1985) model of industry evolution is also related to our analysis. Her model has the same property as the quality ladder models referred to above, where each innovation is introduced by a new firm. The model presented below extends the analysis by Reinganum and Gilbert-Newbery by allowing for competition between horizontally differentiated products. In our model each competing product (variety) is upgraded separately, i.e. each competing product is moving up its own quality ladder.

Our analysis has a similar focus as the work by Ericson and Pakes (1995), Pakes and Ericson (1990) and Pakes and McGuire (1994). They also present a fully specied model of firm growth with investment in innovative activities as the determinant of firm growth. Their model is more sophisticated than ours in that it allows for more extensive strategic interactions and considerations between the firms than we do. They also have a more complete analysis with respect to firm entry and exit. However, the cost of this sophistication is that the Ericson-Pakes model is analytically difficult to handle, and the model must be examined through simulations, while our simpler model is analytically tractable. The two models are to some extent complementary in the sense that our model can most easily be justied in situations where the firms considered are small relative to the market, while the case with many firms creates problems for the simulations of the Ericson-Pakes model (see Pakes and McGuire, 1994).

There are also some earlier studies of ¯rm heterogeneity that are related to our study, such as Jovanovic (1982). (See also Lippman and Rummelt, 1982). Jovanovic's analysis is a dynamic version of Lucas' (1978) model, which is similar to the model by Kihlstrom and La®ont (1979). Both Lucas and Kihlstrom-La®ont present theoretical models that are consistent with heterogeneous ¯rms, but they are both static and the sources of the heterogeneity are given exogenously, as is also the case for Jovanovic's model. The analysis of the relationship between the active and the passive learning models in Pakes and Ericson (1990) is equally relevant for the relationship between Jovanovic's model and the model we present.

Sutton (1997) provides a comprehensive survey of theoretical and empirical studies of ¯rm growth and ¯rm heterogeneity.

Cohen and Klepper (1996) present a model that can rationalize a number of empirical regularities regarding R&D investment and ¯rm size. Our model is in several respects similar to their analysis of the relationship between R&D and size, but it is somewhat more complete as it rationalizes why each ¯rm's pro¯ts from its next innovation are constrained by its current size. More generally, our framework explicitly models ¯rm growth, while ¯rm growth is treated rather brie°y in Cohen and Klepper's model. Klepper (1996) has also examined a related model of ¯rm growth and industry evolution driven by innovation. Klepper's model is able to rationalize a number of interesting empirical regularities of ¯rm growth and industry evolution, but the model is highly stylized at the individual ¯rm level and therefore di±cult to reconcile with structural estimation based on ¯rms as the unit of observation.

Dasgupta (1986) is to some extent also related to our paper, in that he presents a theoretical model consistent with a number of the empirical observations we also consider (and some others). However, Dasgupta's analysis is based on a static model and does not make any predictions about patterns of ¯rm growth and ¯rm heterogeneity. Indeed, his model assumes identical ¯rms. The theoretical literature on preemption and the persistence of incumbency in patent races, that we have already referred to above, is more relevant for our analysis (cf. the papers by Gilbert and Newbery, 1982, 1984).

8 Appendix B: Proposition 3; the formal argument

In this appendix, we will show formally that γ is independent of k: Combining assumption 4 with (11), and using that $\acute{} = °$, it follows that

$$R(k) = e^{°k}\left[\frac{\frac{1}{4}^a\, e^°}{°(r + \gamma)} - \frac{1 + r°}{°\acute{A}_0}\right]:$$

With $\acute{A}(k) = \acute{A}_0 e^{i\,°k}$ and $\gamma = R\acute{A}(k)$, it follows after some algebra that

$$\gamma = \frac{1}{2°}\left\{[4°\frac{1}{4}e^{°\,a}\,\acute{A}_0 + 1]^{1=2} - 1 - 2r°\right\};$$

which does not depend on k.   Q.E.D.

Figure 1: The size distribution for business units in the 8 most R&D intensive (4 digit ISIC) industries in Norwegian manufacturing. The width of the graph corresponds to the number of business units in the industry. The graphs are based on log of sales for the business units, and the boxes show the location of the quartiles in the distribution for each industry, together with the median shown as the horizontal line inside each box. The whiskers indicate the spread of the distribution in the lower and upper quartiles.

.

Figure 2: The pro¯t per unit of output (¼), normalized by marginal costs, as a function of the market share (s_i, in percent), for a range of values of the price elasticity ®. Cf. equation. (16). (The vertical axis shows ¼).