Intel Economics

Paul S. Segerstrom

Stockholm School of Economics

Current version: August 22, 2005

Abstract: This paper presents an endogenous growth model that is designed to be roughly consistent with the experience of high-tech firms like Intel. In the model, industry leaders invest in R&D to improve their products, small firms invest in R&D to become industry leaders and innovating becomes progressively more difficult over time. Consistent with the empirical evidence, the model implies that economic growth is independent of economy size and R&D intensity is independent of firm size. For plausible parameter values, it is optimal to heavily subsidize R&D activities. **JEL classification numbers:** O32, O41. **Key words:** economic growth, R&D.

Author address: Prof. Paul S. Segerstrom, Department of Economics, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden (e-mail: Paul.Segerstrom@hhs.se).

Acknowledgements: I would like to thank two anonymous referees as well as Raouf Boucekkine, Michelle Connally, Carl Davidson, Elias Dinopoulos, Paul Evans, Amy Glass, Gerhard Glomm, Douglas Hibbs, Chol-Won Li, Ulf Jakobsson, Enrique Mendoza, Pietro Peretto, Nancy Stokey, Masako Ueda, Jay Wilson, and Klaus Wälde for helpful comments. I gratefully acknowledge the financial support of the Wallander Foundation and the Research Institute of Industrial Economics (IUI).

1 Introduction

In 1968, engineers Gordon Moore and Bob Noyce left Fairchild, one of the largest semiconductor corporations in Silicon Valley, to found their own start-up company Intel. Initially, Intel made custom memory chips, but in trying to develop some custom circuits for the Japanese calculator manufacturer Busicom, engineers at Intel made a remarkable discovery. They succeeded in imitating at the chip level the architecture of computers by developing a general purpose programable chip. The Intel 4004 chip introduced in 1971 was the world's first microprocessor and maybe the most significant innovation of the 20th century.

The engineers at Intel did not rest on their past accomplishments but immediately went to work on developing more complex and powerful microprocessors. Whereas the Intel 4004 chip contained only 2,300 transistors, the Intel 8080 introduced in 1974 contained 6,000 transistors and was 10 times as powerful as the 4004. The Intel 8080 was in turn followed by the 8086, 286, 386, 486, Pentium and Itanium chips. From 1971 to the present, the number of transistors on an Intel chip has roughly doubled every 2 years, a trend known as "Moore's Law". Intel microprocessor performance, measured in MIPS (millions of instructions per second), has roughly doubled every 18 months and Intel's stock market value has also increased dramatically over time.

Maintaining this pace of innovation, however, has not been easy. Intel has found that developing faster microprocessors becomes progressively more difficult over time. As discussed in Malone (1995), this trend can be traced back to the first microprocessor. It took a team of only 4 engineers to develop the relatively simple Intel 4004 chip in 1971 and most of the work was done by Federico Faggin. In contrast, a team of 20 engineers was involved in designing the considerably more complex Intel 8086 chip in 1978. Intel R&D expenditures have increased dramatically over time, reaching $3.9 billion in the year 2000. Summarizing Intel's history, Malone (1995, p.253) writes, "But miracles, by definition, aren't easy, and in the microprocessor business they get harder all the time. The challenges seem to grow with the complexity of the devices...that is, exponentially."

The goal of this paper is to develop a model of endogenous growth that is roughly consistent with the above-mentioned Intel story. Although Intel is clearly an outlier among high-tech firms in that it has been unusually successful in its R&D activities, Intel nevertheless represents a convenient symbol for high-tech firms in general and the issues studied in this paper have broad applicability. The model has three key properties.

First, industry leaders invest in R&D to improve their own products. Intel has focused on developing faster chips and has repeatedly innovated over time. This type of firm behavior is not at all unusual. Scherer (1980, p.409) cites survey evidence indicating that the lion's share of R&D by industrial firms is directed at improving existing products (vertical R&D). And it is standard practice for industry leaders to invest in R&D activities. Table 1 shows the 2000 net sales and R&D expenditures of several well-known industry leaders. It is clear from this table that not only do many industry leaders devote resources to R&D but their expenditures can be quite substantial.[1]

Second, small firms invest in R&D to become industry leaders. When Intel discovered the world's first microprocessor back in 1971, it was a small firm with just a few employees. Empirical studies reveal that both small and large firms play important roles in contributing to technological change. For example, according to Scherer (1984, chap. 11), companies with fewer than 1,000 employees were responsible for 47.3 percent of important innovations and companies with over 10,000 employees were responsible for 34.5 percent of important innovations.[2]

Third, innovating becomes progressively more difficult over time. Intel has been forced

[1]In endogenous growth models by Romer (1990), Grossman and Helpman (1991b, chap.3) and Jones and Williams (2000), all R&D is directed at developing new horizontally differentiated products. In Segerstrom, Anant and Dinopoulos (1990), Grossman and Helpman (1991a) and Aghion and Howitt (1992), R&D is directed at improving existing products or production processes but none of this R&D is done by the industry leaders themselves. Klette and Kortum (2004) have recently developed a model where industry leaders invest in R&D to improve the products of other firms but do not improve their own products.

[2]In endogenous growth models by Barro and Sala-i-Martin (1995, chap. 7) and Peretto (1998), all R&D is done by industry leaders. Besides the microprocessor, other innovations that have been introduced by small firms include air conditioning, the audio tape recorder, biomagnetic imaging, the digital X-ray, DNA fingerprinting, the FM radio, the hydraulic brake, the integrated circuit, the personal computer, soft contact lens and the vacuum tube (see "Business Innovation," *The Economist*, April 24-30, 2004, p. 75-77.)

to increase its R&D expenditures dramatically over time just to maintain a roughly constant innovation rate because the problems its researchers have wrestled with have been getting progressively harder. And Intel's experience with increasing R&D difficulty appears to be widely shared. At the aggregate level, the number of scientists and engineers engaged in R&D has increased dramatically over time without generating any upward trend in economic growth rates [see Jones (1995b)], and the patents-per-researcher ratio has declined significantly over time in many countries [see Kortum (1997)].[3]

In the related R&D-driven endogenous growth literature, the models that comes the closest to generating firm behavior consistent with the Intel example are Segerstrom and Zolnierek (1999) and Aghion, Harris, Howitt and Vickers (2001). In these models, industry leaders invest in R&D to improve their own products and small firms invest in R&D to become industry leaders. However, innovating does not become more difficult over time and as a consequence, these models have the counterfactual implication that larger economies grow faster. As Jones (1995a) has shown, this scale effect property (which is shared with all the first-generation R&D-driven endogenous growth models) is inconsistent with the time-series evidence from industrialized economies.

In response to the Jones critique, a variety of second-generation R&D-driven endogenous growth models have been developed that do not have the undesirable scale effect property, including Kortum (1997), Segerstrom (1998) and Howitt (1999).[4] However, these models have another significant drawback: they cannot account for Intel's experience of repeatedly innovating over time. In these models, it is not profitable for industry leaders to devote *any* resources to R&D activities due to the Arrow (1962) effect. Firms invest in R&D to become industry leaders but once they succeed, they rest on their past accomplishments and do not try to improve their own products (or production processes).

In this paper, I present a R&D-driven endogenous growth model where both industry leaders and follower firms invest in R&D in each industry. Firms that innovate and develop

---

[3] In many endogenous growth models, including Grossman and Helpman (1991a), Segerstrom and Zolnierek (1999), Aghion, Harris, Howitt and Vickers (2001), and Klette and Kortum (2004), the patents-per-researcher ratio is constant over time. In Romer (1990), the patents-per-researcher ratio increases over time.

[4] A useful review of this literature is Dinopoulos and Thompson (1999).

higher quality products have R&D cost advantages over other firms in improving their own products [as in Barro and Sala-i-Martin (1995, chap.7)]. As a consequence, firms that innovate do not rest on their past accomplishments, but invest in R&D to extend their leadership positions over time. Furthermore, leader R&D expenditure is subject to decreasing returns [as in Segerstrom and Zolnierek (1999)] so follower firms also participate in R&D races.

When follower firms innovate and become industry leaders, they earn monopoly profits as a reward for their past R&D expenditures. When industry leaders innovate, their reward is higher monopoly profits. In each industry, as products improve in quality and become more complex, innovating becomes more difficult [as in Li (2003)]. Because of the increasing cost of innovating, the reward for innovating must also increase over time to justify the increased cost. The reward for innovating increases in each industry because there is positive growth in the population of consumers and quality improvements increase industry demand.

I show that the model has a unique steady-state equilibrium where the industry-level innovation rate is the same in each industry and does not vary over time. In this steady-state equilibrium, R&D employment grows over time without generating any upward trend in the economic growth rate and the patents-per-researcher ratio decreases over time. These properties are consistent with the empirical evidence reported in Jones (1995b) and Kortum (1997). The model also implies that R&D intensity (R&D expenditure as a fraction of total revenue) is independent of firm size, consistent with the empirical evidence reported in Klette and Kortum (2004).

The model is particularly tractable because the value (expected discounted profits) of an industry leader only depends on the quality of its product and not separately on time or other state variables. Along the steady-state equilibrium path, the value of an industry leader jumps up every time the firm innovates and develops a higher quality product. Firms that are unusually successful in innovating achieve unusually high stockmarket values. The value of an industry leader also jumps down when its product is copied by another firm. Thus, the model can account for not only Intel's rise in stockmarket value (third highest in the world at the end of 1998) but also Intel's more recent fall in stockmarket value (with Advanced Micro Devices developing a substitute line of microprocessors).

Turning to welfare implications, I also solve for the R&D subsidy/tax policy that maximizes the discounted utility of the representative household. When there is no copying by firms of other firms' products, it is unambiguously optimal to tax R&D activities. However, allowing for imitation reverses this finding. For plausible parameter values, I find that it is optimal to heavily subsidize R&D. The model's welfare implications are quite sensitive to what is assumed about the rate of copying.[5]

The remainder of this paper is organized as follows. The model is presented in section 2 together with its steady-state equilibrium properties. The welfare properties of the model are explored in section 3. Finally, section 4 contains concluding comments.

2 The Model

2.1 Industry Structure

Consider an economy with a continuum of industries indexed by $\omega \in [0,1]$. In each industry ω, firms are distinguished by the quality j of the products they produce. Higher values of j denote higher quality and j is restricted to taking on integer values. At time $t = 0$, the state-of-the-art quality product in each industry is $j = 0$, that is, some firm in each industry knows how to produce a $j = 0$ quality product and no firm knows how to produce any higher quality product. To learn how to produce higher quality products, firms in each industry participate in R&D races. In general, when the state-of-the-art quality in an industry is j, the next winner of a R&D race becomes the sole producer of a $j+1$ quality product. Thus, over time, products improve as innovations push each industry up its "quality ladder," as in Segerstrom, Anant and Dinopoulos (1990).

In the quality ladders framework, each new product replaces some previously existing product. This appears to be roughly consistent both with Intel's experience as an industry leader (Pentium chips have replaced 486 chips, 486 chips have replaced 386 chips, etc.)

[5] In previous work on optimal R&D policy by Stokey (1995) and Jones and Williams (2000), the rate of copying is assumed to equal zero. Imitation is modelled in Davidson and Segerstrom (1998) but unlike in this paper, the welfare results derived are *ad hoc* (based on by comparing steady-states without solving for the transition paths between steady-states due to policy changes).

and with Intel's experience in becoming an industry leader. Going back to the early 1970s, Intel microprocessors were initially used in minicomputers where they replaced hundreds of individual logic chips. The firms that produced these logic chips, like Fairchild, lost sales to Intel and these losses accelerated as Intel microprocessor performance improved over time.[6]

2.2 Consumers and Workers

The economy has a fixed number of identical households that provide labor services in exchange for wages, and save by holding assets of firms engaged in R&D. Each individual member of a household is endowed with one unit of labor, which is inelastically supplied. The number of members in each family grows over time at the exogenous rate $n > 0$, so the supply of labor in the economy at time t is given by $L(t) = L_0 e^{nt}$. Each household is modelled as a dynastic family which maximizes the discounted utility

$$U \equiv \int_0^\infty e^{-(\rho-n)t} \ln u(t)\, dt \tag{1}$$

where $\rho > n$ is the common subjective discount rate and

$$u(t) \equiv \left[\int_0^1 \left(\sum_j \lambda^j d(j,\omega,t) \right)^{\alpha} d\omega \right]^{\frac{1}{\alpha}} \tag{2}$$

is the utility per person at time t. Equation (2) is a quality-augmented Dixit-Stiglitz consumption index: $d(j,\omega,t)$ denotes the quantity consumed of a product of quality j produced in industry ω at time t, $\lambda > 1$ measures the size of quality improvements, and $\alpha \in (0,1)$ determines the elasticity of substitution between industries $\sigma = 1/(1-\alpha)$. Because λ^j is increasing in j, (2) captures in a simple way the idea that consumers prefer higher quality products.

Utility maximization involves two steps. First, each household allocates per capita expenditure $c(t)$ to maximize $u(t)$ given the prevailing market prices $p(j,\omega,t)$ at time t.

[6] Fairchild invented the planar process in 1958 and was the dominant semiconductor firm in the 1960s but by 1980, it had become an insignificant player in this industry (see Malone (1995)). As was mentioned in the introduction, the founders of Intel were former Fairchild employees.

Solving this optimal control problem yields the per capita demand function

$$d(\omega, t) = \frac{q(\omega, t)p(\omega, t)^{-\sigma} c(t)}{\int_0^1 q(\omega, t)p(\omega, t)^{1-\sigma} d\omega} \tag{3}$$

for the product $j(\omega, t)$ in industry ω with the lowest quality-adjusted price $p(j, \omega, t)/\lambda^j$, where $q(\omega, t) \equiv \delta^{j(\omega,t)}$ is an alternative measure of product quality and $\delta \equiv \lambda^{\sigma-1}$. The quantity demanded for all other products is zero. To break ties, I assume that when quality adjusted prices are the same for two products of different quality, each consumer only buys the higher quality product. Second, each household maximizes discounted utility (1) given (2), (3) and the intertemporal budget constraint. Solving this optimal control problem yields the well-known intertemporal optimization condition

$$\frac{\dot{c}(t)}{c(t)} = r(t) - \rho, \tag{4}$$

where $r(t)$ is the instantaneous rate of return at time t. This differential equation must be satisfied throughout time in equilibrium and implies that a constant per capita expenditure path is optimal only when the market interest rate equals ρ. A higher market interest rate induces consumers to save more now and spend more later, resulting in increasing per capita consumption over time. Equation (4) implies that in any steady-state equilibrium, the market interest rate r must be constant over time.

2.3 Product Markets

In each industry, firms compete in prices. If two or more firms charge the same price and sell the same quality product, then they share consumers equally. Labor is the only input used in production and there are constant returns to scale. One unit of labor is required to produce one unit of output, regardless of quality. Labor markets are perfectly competitive and the wage is normalized to unity throughout time.[7] Consequently, each firm has a constant marginal cost of production equal to one.

[7] In classical general equilibrium models, the choice of numeriare at each point in time is of no consequence. However, one must remember that all prices are nominal prices (including wages and interest rates) and need to be appropriately adjusted to get real prices.

At any point in time, a firm can choose to shut down its production facilities and once it has done so, this decision can only be reversed by incurring a positive entry cost. Furthermore, each firm that fails to attract any consumers (has zero sales) incurs a positive cost of maintaining its unused production facilities, in addition to the constant marginal cost of production mentioned above.[8] Thus firms that are not able to attract any consumers in equilibrium (because of the low relative quality of their products) choose to immediately shut down their production facilities and do not play any role in determining market prices.

The profits earned by firms depend on the level of competition in each industry. If there are two or more firms that produce the state-of-the-art quality product, then Bertrand price competition implies that each firm charges a price equal to marginal cost (one), sells to its share of consumers and earns zero economic profits. On the other hand, if there is only one firm that produces the state-of-the-art quality product, then this quality leader earns positive economic profits until its product is copied or surpassed by another firm.

To determine the economic profits that a single quality leader earns, consider what happens when a new firm innovates and becomes the only quality leader in its industry. This firm's closest competitors are follower firms one step down in the industry's quality ladder, including the previous quality leader. Letting p_L denote the new quality leader's price, it is either in the interest of the new quality leader to charge the limit price $p_L = \lambda$ [a price that is just low enough so that the follower firms cannot compete, as in Grossman and Helpman (1991a)] or to charge the unconstrained monopoly price $p_L = \frac{1}{\alpha}$ [a price that is obtained by maximizing the quality leader's profit flow $\pi_L = (p_L - 1)d(\omega, t)L(t)$ with respect to p_L]. In either case, the follower firms do not attract any consumers and have zero sales. Given the positive costs of maintaining unused production facilities, it is profit-maximizing for the remaining firms with less than state-of-the-art quality products to immediately shut down their production facilities. Thus, using (3), the single quality leader ends up charging the unconstrained monopoly price $p_L = \frac{1}{\alpha}$ and earns the monopoly profit

[8] I have in mind costs like the costs of employing security guards to prevent looting at unused production facilities, guards that would not be needed if the production facilities were being used.

flow

$$\pi_L = (p_L - 1)\frac{q(\omega,t)}{Q(t)}y(t)L(t), \tag{5}$$

where $Q(t) \equiv \int_0^1 q(\omega,t)d\omega$ is the average quality level across industries and

$$y(t) = \frac{Q(t)p_L^{-\sigma}c(t)}{\int_0^1 q(\omega,t)p(\omega,t)^{1-\sigma}d\omega} \tag{6}$$

is the per-capita quantity demanded for a single quality leader's product when the product is of average quality. The single quality leader's profit flow π_L is an increasing function of the per-unit profit margin $p_L - 1$, the relative quality of the firm's product $\frac{q(\omega,t)}{Q(t)}$, and the market size measure $y(t)L(t)$.[9]

2.4 R&D Races

Labor is the only input used to do R&D in any industry, is perfectly mobile across industries and between production and R&D activities. Firms make their R&D choices simultaneously and I solve for Nash equilibrium behavior.

In each industry, there are two types of firms that can hire R&D workers: R&D leaders and R&D followers. I distinguish between quality leaders (the firms that produce the highest quality products) and R&D leaders (the firms that have the best R&D technologies). A firm can become a quality leader by either innovating or copying another firm's product. In contrast, a firm can only become a R&D leader by innovating. When firms innovate, they gain some knowledge that is useful for further innovating that is not acquired by merely copying another firm's product.

I begin by describing the R&D technology of R&D followers: all firms besides the current R&D leader in an industry (there is free entry by R&D followers into each R&D race). A follower firm i that hires ℓ_i units of R&D labor in industry ω at time t is successful in discovering the next higher quality product $j(\omega,t)+1$ with instantaneous probability[10]

$$I_i = A_F \frac{\ell_i}{\delta^{j(\omega,t)}}, \tag{7}$$

[9]In section 3.4, I relax the assumption that there is a positive cost of maintaining unused production facilities and study how the model's properties change when new quality leaders charge limit prices ($p_L = \lambda$) instead of monopoly prices ($p_L = \frac{1}{\alpha}$).

[10]By instantaneous probability, I mean that $I_i dt$ is the probability that firm i will innovate by time $t + dt$

where $A_F > 0$ is a follower firm R&D productivity parameter. With this R&D technology, there is constant returns to scale: doubling R&D labor ℓ_i doubles the firm's innovation rate I_i. Since the denominator is increasing in j, equation (7) captures the idea that as products become more complex with each step up the quality ladder, innovating becomes progressively more difficult.[11]

For the current R&D leader in an industry (the firm that most recently innovated), this firm has access to a better R&D technology. When the leader firm hires ℓ_L units of R&D labor, this firm is successful in discovering the next higher quality product with instantaneous probability

$$I_L = A_L \left(\frac{\ell_L}{\delta^{j(\omega,t)}} \right)^{\beta}, \tag{8}$$

where $A_L > 0$ is a R&D leader productivity parameter and $\beta < 1$ measures the degree of decreasing returns to leader R&D expenditure. The parameter restriction $\beta < 1$ implies that R&D workers employed by R&D leaders are more productive on the margin than R&D workers employed by R&D followers (when the scale of R&D operations is not too large).[12]

The distinction between R&D leaders and R&D followers is made to guarantee that both large and small firms participate in R&D races. If all firms had access to the same R&D technology (7), then quality leaders (large firms) would not invest in R&D. Firms would invest in R&D to become quality leaders but once they succeeded, they would rest on their past accomplishments and not try to improve their own products [as in Grossman and Helpman (1991a)]. Large firms need some R&D cost advantages to justify investing in R&D and it is not enough to assume that $A_L > A_F$ (with $\beta = 1$) because then both large

conditional on not having innovated by time t, where dt is an infinitesimal increment of time. Alternatively stated, I_i is the Poisson arrival rate of innovations by firm i.

[11]The reason for using the same δ parameter in both the demand equation (3) and the technology equation (7) is explained in Barro and Sala-i-Martin (1995, p.249-250). Without this assumption, innovation rates would either gradually increase or gradually decrease over time. With this assumption, every time innovation occurs in an industry, the costs of innovating rise by the same magnitude as the benefits from innovating, so the model has a steady-state equilibrium with a constant innovation rate in each industry.

[12]Regardless of the values of A_L and A_F, $\frac{dI_L(\ell)}{d\ell_L} > \frac{dI_i(\ell)}{d\ell_i}$ for sufficiently small ℓ when $\beta < 1$.

and small firms participate in R&D races only for a knife-edge set of parameter values [as is shown in Segerstrom and Zolnierek (1999)]. Thus, $\beta < 1$ is assumed to guarantee that both large and small firms participate in R&D races for a nontrivial range of parameter values.[13]

The returns to doing R&D are independently distributed across firms, across industries, and over time. Thus in a typical industry, the industry-wide instantaneous probability of R&D success is simply $I \equiv I_L + I_F = I_L + \sum_i I_i$ where I_F is the instantaneous probability of R&D success by all R&D follower firms combined.

The focus of this paper is on the balanced growth (or steady-state) properties of the model where all endogenous variables grow at constant (not necessarily the same) rates. The model will be solved for a steady-state equilibrium where the innovation rate I does not vary across industries. However, since the incentives for R&D leaders to innovate depend on whether they are currently earning monopoly profits or not, I need to introduce a notational distinction. Let I_L and I_F denote the innovation rates by R&D leaders and followers respectively when monopoly profits are earned in the product market and let I_{LF} and I_{FL} denote the innovation rates by R&D leaders and followers respectively after copying has occured. It will be shown that a steady-state equilibrium with $I = I_L + I_F = I_{LF} + I_{FL}$ exists where I, I_L, I_F, I_{LF} and I_{FL} are all constants over time.

Copying is modelled in the simplest possible way. There is an exogenous instantaneous probability $C \geq 0$ that some R&D follower firm succeeds in copying the current quality leader's production technology. Then both the leader and the copying firm earn zero economic profits due to price-setting and Bertrand competition. The copying firm still has access to the follower R&D technology and thus is a follower when it comes to R&D. The rate of copying C does not vary across industries or over time. As illustrated in Figure 1,

[13] Etro (2004) explores an alternative approach to explaining why large firms invest in R&D that does not rely on R&D cost advantages. Instead of solving for a Nash equilibrium in each R&D race, Etro assumes that the single quality leader can make a strategic precommitment to a level of R&D investment and solves for a Stackelberg equilibrium. One drawback of Etro's approach is that it implies a fairly low persistence of quality leaders. For plausible parameter values, Etro (2004, p.301) calculates a probability of only .084 that the current quality leader discovers the next innovation.

each industry fluctuates between two states over time, with innovation resulting in a sin-

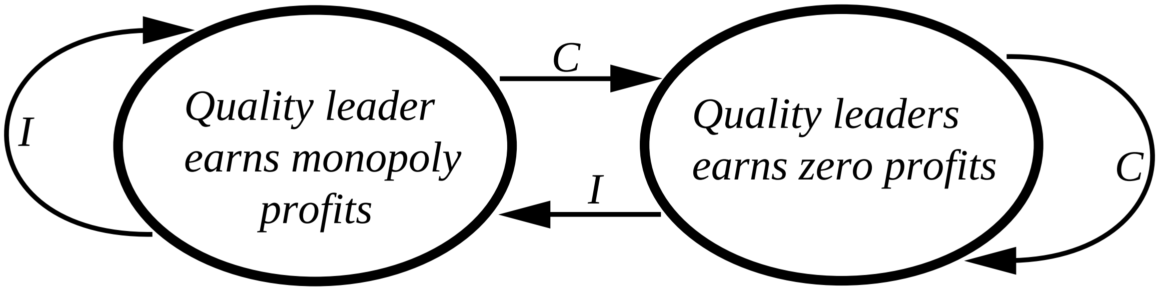

Figure 1: The steady-state pattern of innovation and copying

gle quality leader earning monopoly profits and copying resulting in more than one quality leader earning zero economic profits.

2.5 R&D Optimization

All firms are assumed to maximize their expected discounted profits in deciding how much to invest in R&D activities. To maximize expected discounted profits, both leaders and followers must solve stochastic optimal control problems where the state variable $j(\omega, t)$ in each industry ω is a Poisson jump process with intensity I and magnitude $+1$. The model is particularly tractable because the value (expected discounted profits) of a single quality leader v_L only depends on the quality of its product $j(\omega, t)$ and not separately on t, ω or any other state variables. The same holds true for the value of a R&D leader whose product has been copied: v_{LF}. Free entry and constant returns to scale imply that R&D followers have no market value: $v_F = v_{FL} = 0$. To simplify notation, I will henceforth let j_ω denote the state-of-the-art quality level in industry ω instead of $j(\omega, t)$ and leave the functional dependence on t implicit.

For a single quality leader, the relevant Hamilton-Jacobi-Bellman equation is[14]

$$\begin{aligned} r \cdot v_L(j_\omega) = \max_{\ell_L} \quad & \pi_L(j_\omega, t) - (1-s)\ell_L + I_L \left[v_L(j_\omega + 1) - v_L(j_\omega) \right] \\ & + I_F \left[v_F(j_\omega + 1) - v_L(j_\omega) \right] \\ & + C \left[v_{LF}(j_\omega) - v_L(j_\omega) \right]. \end{aligned} \tag{9}$$

A single quality leader earns the monopoly profit flow $\pi_L(j_\omega, t)$ today and also incurs the R&D cost $(1-s)\ell_L$. With instantaneous probability I_L, the leader innovates (learns how to produce a $j_\omega + 1$ quality product) and its value jumps up as a result. However, with instantaneous probability I_F, some R&D follower firm innovates and the leader becomes a R&D follower. Also with instantaneous probability C, some R&D follower firm copies the leader's product and the leader becomes a copied R &D leader. Equation (9) states that the maximized expected return on a leader firm's stock must equal the return on an equal-sized investment in a riskless bond.

For a R&D leader in an industry where copying has occurred and the R&D leader competes directly against R&D followers in the product market, the relevant Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} r \cdot v_{LF}(j_\omega) = \max_{\ell_{LF}} \quad & -(1-s)\ell_{LF} + I_{LF} \left[v_L(j_\omega + 1) - v_{LF}(j_\omega) \right] \\ & + I_{FL} \left[v_F(j_\omega + 1) - v_{LF}(j_\omega) \right] \\ & + C \left[v_{LF}(j_\omega) - v_{LF}(j_\omega) \right]. \end{aligned} \tag{10}$$

The R&D leader incurs the R&D cost $(1-s)\ell_{LF}$ today but earns no profit flow. With instantaneous probability I_{LF}, the R&D leader innovates, learns how to produce a $j_\omega + 1$ quality product and becomes only quality leader. However, with instantaneous probability

[14] See Malliaris and Brock (1982, p.123-124) for the application of stochastic dynamic programming techniques to Poisson jump processes. These techniques have been used to study the R&D incentives of industry leaders is several earlier papers. Thompson and Waldo (1994) analyze a model where only industry leaders can improve their own products and entry is not allowed. Aghion, et al (2001) study the case where firms must first do imitative R&D to catch up with industry leaders and then they can do innovative R&D to improve their own products. More recently, Klette and Kortum (2004) have developed a model where industry leaders do innovative R&D to improve other industry leaders' products and become multi-product firms.

I_{FL}, some R&D follower innovates and the R&D leader becomes a R&D follower in the next R&D race. Also, with instantaneous probability C, some firm copies the R&D leader's product, in which case the R&D leader continues to directly compete against other quality leaders in the product market.

For R&D follower firm i in an industry with a single quality leader, the relevant Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} r \cdot v_F(j_\omega) = \max_{\ell_i} \quad & -(1-s)\ell_i + I_i \left[v_L(j_\omega + 1) - v_F(j_\omega) \right] \\ & + (I_{-i} + I_L) \left[v_F(j_\omega + 1) - v_F(j_\omega) \right] \\ & + C \left[v_{FL}(j_\omega) - v_F(j_\omega) \right], \end{aligned} \tag{11}$$

where $I_{-i} \equiv I_F - I_i$ is the R&D intensity by all other R&D followers combined and s is the R&D subsidy rate. Each follower incurs the R&D cost $(1-s)\ell_i$ today but earns no profit flow. With instantaneous probability I_i, the follower innovates, becomes a quality leader and learns how to produce a $j_\omega + 1$ quality product. However, with instantaneous probability $I_{-i} + I_L$, some other firm innovates (either the current R&D leader or another R&D follower) and the follower continues to be a follower in the next R&D race. Also, with instantaneous probability, some firm copies the quality leader's product, and the R&D follower continues to be a follower in the next R&D race.

In the rest of the paper, I focus on the case where both R&D leaders and followers participate in each R&D race, that is, $I > I_L > 0$ and $I > I_{LF} > 0$. Such an interior solution to the Bellman equations exists if and only if $A_L > 0$ is sufficiently small, as I will demonstrate shortly. The interior solution is given by

$$I_L = A_L \left[\frac{A_L \beta}{A_F} \left(1 - \frac{1}{\delta} \right) \right]^{\frac{\beta}{1-\beta}} \tag{12}$$

$$I_{LF} = A_L \left[\frac{A_L \beta}{A_F} \left(1 - \frac{1}{\hat{\delta}} \right) \right]^{\frac{\beta}{1-\beta}} \tag{13}$$

$$v_L(j_\omega) = \frac{(1-s)\delta^{j_\omega}}{\delta A_F} \tag{14}$$

$$v_{LF}(j_\omega) = \frac{(1-s)\delta^{j_\omega}}{\hat{\delta} A_F} \tag{15}$$

$$r + I = \delta A_F \frac{p_L - 1}{1 - s} \frac{y}{x} + f(\delta) + C \left(\frac{\delta - \hat{\delta}}{\hat{\delta}} \right) \quad (16)$$

$$r + I = f(\hat{\delta}) \quad (17)$$

where the function f is defined by $f(\hat{\delta}) \equiv \hat{\delta} I_{LF} - \hat{\delta} A_F \left(\frac{I_{LF}}{A_L}\right)^{1/\beta}$, $r = \rho$, $I = n/(\delta - 1)$, equation (17) determines the steady-state value of $\hat{\delta}$, and $x(t) \equiv \frac{Q(t)}{L(t)}$.

The proof of the above-mentioned claims is presented in the Appendix. In the main text, I want to focus on the economic interpretation of this solution.

Equation (14) implies that the value of a single quality leader v_L jumps up every time the firm innovates and develops a higher quality product (j_ω increases). Also, R&D followers respond to either an increase in the R&D subsidy rate s or an increase in their R&D productivity parameter A_F by innovating more frequently and with quality leaders being driven out of business more frequently, the value of being a single quality leader naturally falls. Equation (14) provides a simple explanation for Intel's high stock market value: Intel has been unusually successful in its R&D activities, leading to a high value of j_ω, higher than in most other industries. In the microprocessor industry, because it is more difficult to innovate, the reward for innovating has to be higher to induce R&D effort. In this model, the assumption of increasing R&D difficulty plays a central role in explaining why some firms have higher stock market valuations than other firms.[15]

Equation (12) implies that the innovation rate for a single quality leader I_L is completely pinned down by parameter values and does not change over time or vary across industries. The constancy of I_L together with (8) implies that a single quality leader's R&D employment ℓ_L is constant during an R&D race but jumps up every time the firm innovates. This property in turn helps to explain why aggregate R&D employment gradually increases over time as firms (R&D leaders as well as followers) innovate in a wide variety of industries. Equation (12) has very intuitive implications. Other things being equal, single quality leaders are more innovative (I_L increases) when their R&D workers are more

[15] At the end of 1998, Intel's stock market capitalization was $194 billion, the third highest in the world (behind only Microsoft and General Electric). Source: "Unbearable lightness of being," by John Plender, *Financial Times*, December 8, 1998. Since 1998, Intel's stock market value has dropped due to increased competition in the microprocessor industry.

productive (A_L increases) or innovations represent larger improvements in product quality (λ and δ increase). On the other hand, changes in the structure of the economy that make it more attractive for R&D followers to invest in R&D have an adverse effect on the relative R&D effort of industry leaders. Single quality leaders are less innovative (I_L decreases) when follower firm R&D workers are more productive (A_F increases).

Equations (13) and (15) describing a copied quality leader have similar economic interpretations. The property $\hat{\delta} > \delta > 1$ established in the Appendix implies that $I_{LF} > I_L > 0$ (R&D leaders are more innovative after they have been copied, since firms that are earning monopoly profits have less to gain from further innovation) and $v_L(j_\omega) > v_{LF}(j_\omega) > 0$ (the market value of a R&D leader drops when its product is copied by another firm). The market value of a R&D leader does not drop all the way to zero when its product is copied since the leader's R&D capability has positive market value.

The endogenous variable $x(t) \equiv \frac{Q(t)}{L(t)}$ denotes the average quality of products relative to the size of the economy. As product quality improves over time (Q increases), innovating becomes more difficult. On the other hand, as the economy increases in size over time (L increases), there are more resources that can be devoted to innovating. Thus x is a natural measure of relative R&D difficulty.

Equation (16) is the *steady-state R&D condition*, one of the key equations in the model. It characterizes the steady-state conditions that must prevail if firms are making profit-maximizing R&D decisions. Equation (16) implies that when relative R&D difficulty x is higher, the individual consumer demand measure y and the monopoly profits from innovating π must be higher to justify the R&D expenditures of firms.

Finally, the above-described interior solution ($I > I_{LF} > I_L > 0$) exists if and only if $A_L > 0$ is sufficiently small. The assumption that $A_L > 0$ is sufficiently small means that, in each industry, the R&D productivity of leaders is sufficiently low so that leaders do not do all the R&D. Some R&D is always done by followers (small firms).

2.6 Quality Dynamics

To calculate how $Q(t)$ evolves over time in a steady-state equilibrium, recall that the average quality of products at time t is $Q(t) = \int_0^1 \delta^{j_\omega}\, d\omega$. Since j_ω jumps up to $j_\omega + 1$ when innovation occurs in industry ω, and the innovation rate I is constant across industries, the time derivative of $Q(t)$ is

$$\dot{Q}(t) = \int_0^1 \left[\delta^{j_\omega+1} - \delta^{j_\omega}\right] I(t)\, d\omega = (\delta - 1)I(t)Q(t). \tag{18}$$

The growth rate of average product quality $\frac{\dot{Q}}{Q}$ is proportional to the innovation rate I, which must be constant over time in a steady-state equilibrium as was earlier claimed.

Differentiating $x(t) \equiv Q(t)/L(t)$ and using (18) yields a state equation that describes how relative R&D difficulty $x(t)$ evolves over time:

$$\dot{x}(t) = f_1(x(t), I(t)) \tag{19}$$

where the f_1 function is defined by $f_1(x, I) \equiv x[(\delta - 1)I - n]$. Note that (19) holds both in and out of steady-state. In any steady-state equilibrium where x is constant over time, (19) implies that the steady-state innovation rate is

$$I = \frac{n}{\delta - 1}, \tag{20}$$

as was earlier claimed. The steady-state innovation rate depends only on the population growth rate n and the R&D difficulty parameter δ, as in Segerstrom (1998) and Li (2003). In a steady-state equilibrium, individual researchers are becoming less productive and firms compensate for this by increasing the number of employed researchers over time. This compensation is only feasible for firms in general if there is positive population growth, so positive population growth is needed to sustain technological change in the long run.

The average quality of products $Q(t)$ can be broken up into two parts

$$Q(t) = \int_0^1 \delta^{j_\omega}\, d\omega = Q_L(t) + Q_C(t) = \int_{m_L} \delta^{j_\omega}\, d\omega + \int_{m_C} \delta^{j_\omega}\, d\omega,$$

where m_L is the set of industries where there is a single quality leader, m_C is the set of industries where copying has occurred (there is more than one quality leader), $Q_L(t)$ is a measure of product quality in industries where there is a single quality leader and $Q_C(t)$ is a measure of product quality in industries where copying has occurred.

Given the time path of Q, the time paths of Q_L and Q_C can be determined once one knows how $q_L(t) \equiv \frac{Q_L(t)}{Q(t)}$ evolves over time. Referring back to Figure 1, the time derivative of Q_L is

$$\begin{aligned}\dot{Q}_L &= \int_{m_C} \delta^{j_\omega+1} I \, d\omega - \int_{m_L} \delta^{j_\omega} C \, d\omega + \int_{m_L} \left[\delta^{j_\omega+1} - \delta^{j_\omega}\right] I \, d\omega \\ &= I\delta Q_C - CQ_L + (\delta - 1)IQ_L.\end{aligned}$$

Since $\frac{\dot{q}_L}{q_L} = \frac{\dot{Q}_L}{Q_L} - \frac{\dot{Q}}{Q}$, it immediately follows that

$$\dot{q}_L = f_2(I, q_L) \tag{21}$$

where the $f_2(\cdot)$ function is defined by $f_2(I, q_L) \equiv \delta I(1 - q_L) - q_L C$. Note that (21) holds both in and out of steady-state. In a steady-state equilibrium where q_L is constant over time, (21) implies that

$$q_L = \frac{\delta I}{C + \delta I} \tag{22}$$

The steady-state value of q_L is completely determined by the parameter values C, n and δ.

2.7 The Labor Market

All workers are employed by firms in either production or R&D activities. The wage rate adjusts so all workers are fully employed at each point in time. The total labor supply at time t is simply $L(t)$. In what follows, I will often suppress the functional dependence of variables on time to simplify notation.

I will first solve for total production employment. The market price is p_L in m_L industries with a single quality leader and 1 in m_C industries with more than one quality leader. Using (3) and (6), production employment in a m_L industry is

$$d(\omega, t)L(t) = \frac{q(\omega, t)p_L^{-\sigma}c(t)L(t)}{\int_0^1 q(\omega, t)p(\omega, t)^{1-\sigma}d\omega} = \frac{q(\omega, t)}{Q(t)}y(t)L(t)$$

and production employment in a m_C industry is

$$d(\omega, t)L(t) = \frac{q(\omega, t)c(t)L(t)}{\int_0^1 q(\omega, t)p(\omega, t)^{1-\sigma}d\omega} = \frac{q(\omega, t)}{Q(t)}p_L^{\sigma}y(t)L(t).$$

Thus total production employment in the economy is

$$\int_{m_L} d(\omega,t)L(t)\,d\omega + \int_{m_C} d(\omega,t)L(t)\,d\omega = yLf_3(q_L)$$

where the f_3 function is defined by

$$f_3(q_L) \equiv q_L + p_L^\sigma(1-q_L).$$

Note that f_3 is a decreasing function of q_L. As q_L increases and there is more monopoly pricing in the economy, there is less need for production workers (taking y and L as given) since firms with higher prices sell less output.

Although attention is restricted in this paper to the steady-state properties of the model, to characterize the welfare-maximizing steady-state outcome, I need to specify what the labor market condition looks like for *some* outcomes outside of the steady-state. So I fix I_L and I_{LF} at their steady-state values [given by (12) and (13)] but allow I to deviate from its steady-state value [given by (20)]. By allowing I to deviate from its steady-state value, since $I = I_L + I_F$ and $I = I_{LF} + I_{FL}$, I am effectively allowing I_F and I_{FL} to deviate from their steady-state values as well.

Next I will solve for total R&D employment. Using (7) and (8), R&D employment in a m_L industry is

$$\ell_L + \ell_F = \left[\left(\frac{I_L}{A_L}\right)^{1/\beta} + \frac{I_F}{A_F}\right] q(\omega,t)$$

and R&D employment in a m_C industry is

$$\ell_{LF} + \ell_{FL} = \left[\left(\frac{I_{LF}}{A_L}\right)^{1/\beta} + \frac{I_{FL}}{A_F}\right] q(\omega,t).$$

Thus total R&D employment in the economy is

$$\int_{m_L} (\ell_L + \ell_F)\,d\omega + \int_{m_C} (\ell_{LF} + \ell_{FL})\,d\omega = Qf_4(I,q_L)$$

where the f_4 function is defined by

$$f_4(I,q_L) \equiv \frac{I}{A_F} + \left[\left(\frac{I_L}{A_L}\right)^{1/\beta} - \frac{I_L}{A_F}\right] q_L + \left[\left(\frac{I_{LF}}{A_L}\right)^{1/\beta} - \frac{I_{LF}}{A_F}\right](1-q_L)$$

It is straighforward to verify that the f_4 function is increasing in both I and q_L. Other things being equal, more innovation (an increase in I) is associated with more R&D employment and because more monopoly pricing (an increase in q_L) leads to a less efficient utilization of R&D resources between leaders and followers, more monopoly pricing is associated with more R&D employment. Note that total R&D employment is also increasing in Q. As the average quality of products Q increases over time, more workers need to be employed in R&D activities to maintain the innovation rate I. It is in this sense that innovating becomes progressively more difficult over time.

Putting things together, full employment of labor in the economy at time t implies that $L(t) = y(t)L(t)f_3(q_L(t)) + Q(t)f_4(I(t), q_L(t))$. Dividing both sides by $L(t)$ yields the resource condition

$$1 = y(t)f_3(q_L(t)) + x(t)f_4(I(t), q_L(t)) \tag{23}$$

Equation (23) holds both in and out of steady-state and will be used to characterize the welfare-maximizing steady-state equilibrium.

In any steady-state equilibrium where the innovation rate I is constant over time, (22) implies that q_L is constant over time. It then follows from (23) that relative R&D difficulty x is constant over time, as was earlier claimed. It also follows from (23) that the demand measure y is constant over time and given (6),

$$y = \frac{Qp_L^{-\sigma}c}{\int_0^1 p^{1-\sigma}q\,d\omega} = \frac{p_L^{-\sigma}c}{p_L^{1-\sigma}q_L + 1 - q_L} \tag{24}$$

implies that individual consumer expenditure c is constant over time. Equation (4) then implies that the steady-state interest rate r must equal the consumer's subjective discount rate ρ, as was earlier claimed.

Since c, q_L, x and I are all constant over time in any steady-state equilibrium, the *steady-state resource condition* is simply

$$1 = yf_3(q_L) + xf_4(I, q_L). \tag{25}$$

The interpretation of (25) is quite intuitive. The first term on the right-hand side $yf_3(q_L)$ is the fraction of workers that are employed in production activities. This fraction is increasing in individual consumer expenditure c and is decreasing in the degree of monopoly pricing that prevails in the economy, since higher prices are associated with lower sales [f_3 is a decreasing function of q_L]. The second term on the right-hand side $xf_4(I, q_L)$ is the fraction of workers that are employed in R&D activities. This fraction is increasing in relative R&D difficulty x, is increasing in the innovation rate I and is increasing in the degree of monopoly pricing that prevails in the economy [f_4 is an increasing function of both I and q_L]. The last property is due to the fact that as the degree of monopoly pricing in the economy increases (due to a reduction in the rate of copying C), there is a less efficient utilization of R&D resources between leaders and followers and more R&D workers are needed to maintain a given innovation rate I.

2.8 The Steady-State Equilibrium

Both the steady-state R&D and resource conditions are illustrated in Figure 2. The steady-

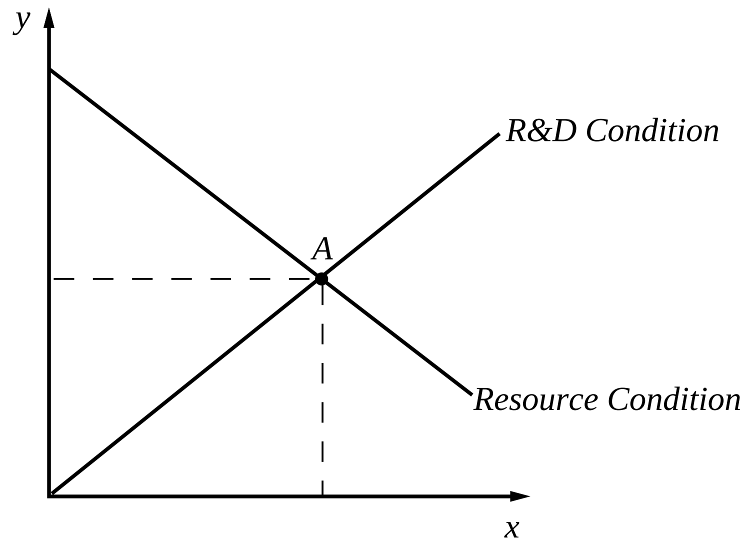

Figure 2: The steady-state equilibrium

state R&D condition (16) is upward-sloping in (x, y) space, indicating that when R&D is relatively more difficult, individual consumer demand y and the monopoly profits from innovating π must be higher to justify the R&D expenditures of firms. The steady-state resource condition (25) is downward-sloping in (x, y) space, indicating that when R&D is relatively more difficult and more resources are used in the R&D sector to maintain the steady-state innovation rate I, less resources are available to produce goods for consumers. The unique intersection between the R&D and resource conditions at point A determines the steady-state equilibrium values of individual consumer demand y and relative R&D

difficulty x.[16]

2.9 Comparative Steady-State Equilibrium Exercises

To illustrate the properties of the model, I consider two comparative steady-state exercises: increasing s and increasing C.

First, consider the steady-state effects of increasing the R&D subsidy rate s. The steady-state R&D condition (16) is linear (x, y) space and goes through the origin. Furthermore, an increase in s is associated with a decrease in y holding x fixed. In constrast, s does not appear in the steady-state resource condition (25). Thus, as is illustrated in Figure 3, an increase in s causes the steady-state R&D condition to pivot downwards and has no effect on the steady-state resource condition.

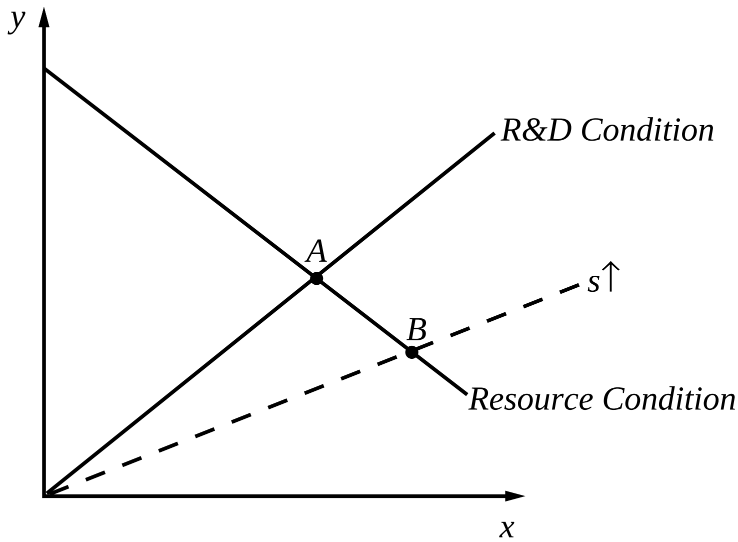

Figure 3: An Increase in the R&D Subsidy Rate

An increase in the R&D subsidy rate s causes individual consumer demand y to decrease and relative R&D difficulty x to increase. Now $x(t) = \frac{Q(t)}{L(t)}$ is a state variable in the model that can only gradually adjust over time. Thus, the correct way to interpret this

[16]It is worth remembering that the nominal wage rate has been normalized to equal one throughout time, so c represents nominal consumer expenditure (and y represents nominal consumer demand). The constancy of c over time in a steady-state equilibrium is associated with positive growth in real consumer expenditure since technological change increases the real wage over time.

steady-state effect is that, along the transition path from the old steady-state equilibrium (at point A in Figure 3) to the new steady-state equilibrium (at point B in Figure 3), the average quality of products $Q(t)$ must grow at a faster rate than labor supply $L(t)$. Using (18), a higher R&D subsidy rate s is associated with a temporary increase in the innovation rate $I(t)$, that is, a temporary acceleration in the rate of technological change. Even though a higher R&D subsidy has no effect on the steady-state innovation rate $I = \frac{n}{\delta - 1}$, the effects of a higher R&D subsidy are quite intuitive: technological change is promoted (temporarily) and the share of total employment in R&D increases (permanently, since $xf_4(I, q_L)$ increases in the steady-state resource condition). Summarizing, I have established

Theorem 1 *A permanent increase in the R&D subsidy rate s leads to a temporary increase in the rate of technological change and a permanent increase in the share of total employment in R&D activities.*

Although the focus of this paper is on steady-state outcomes and the transitional dynamic properties of the model are not formally studied, earlier papers by Segerstrom (1998) and Li (2003) have studied transitional dynamics in simpler settings where only follower firms engage in R&D activities. They find that the unique steady-state equilibrium is globally stable: regardless of initial conditions, the equilibrium transition path involves convergence over time to the steady-state outcome. Furthermore, Steger (2003) has investigated the speed of convergence issue for the Segerstrom (1998) model. He finds that for plausible parameter values, convergence to the steady-state tends to be very slow, so the "temporary" effects of changes in public policies last a long time. Steger's calibrated model yields a half-life of 38 years, that is, it takes 38 years to go half the distance to the steady-state, which is roughly consistent with empirical studies on the speed of convergence [i.e., Barro and Sala-i-Martin (1992)]. Steger's finding is useful for thinking about Figure 3: following an increase in the R&D subsidy rate s, one should expect it to take roughly 40 years to go half the distance from point A to point B (measuring along the horizontal axis).[17]

[17]One frequently raised criticism of "semi-endogenous" growth models [i.e., Jones (1995b), Segerstrom (1998) and this paper] is that, while these models solve the scale effect problem, they do so in such a way as to eliminate the long-run growth effects of public policies. According to the critics, the whole point of

Second, consider the steady-state effects of increasing the rate of copying C. Since an increase in C is associated with a increase in y holding x fixed, the steady-state R&D condition (16) pivots upwards in (x, y) space, as illustrated in Figure 4. Things are more complicated for the steady-state resource condition (25). An increase in C is associated with a decrease in q_L and therefore an increase in $f_3(q_L)$ and a decrease in $f_4(I, q_L)$. It follows that an increase in C causes the vertical intercept of the steady-state resource condition to fall and the horizontal intercept to rise. As illustrated in Figure 4, starting from an

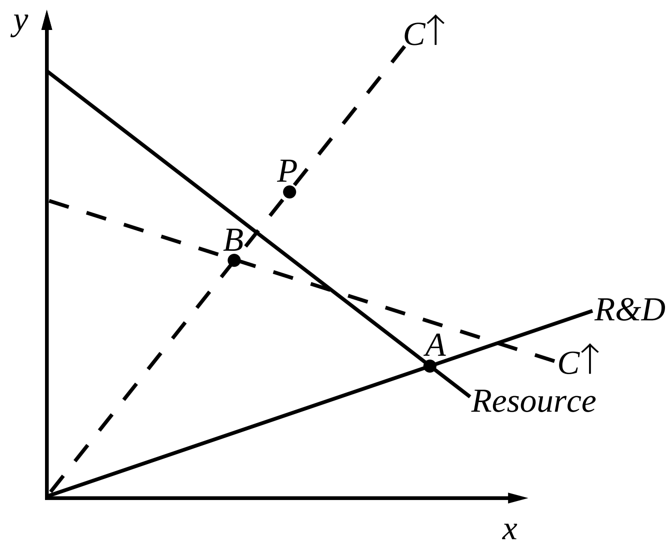

Figure 4: An Increase in the Rate of Copying

initial steady-state equilibrium (given by point A), an increase C lead to a new steady-state equilibrium (given by point B). The increase in C is associated with an increase in y and a decrease in x as illustrated but Figure 4 could have been drawn differently (with the new resource condition going though point P instead of point B, for example). What one can say definitively is that either an increase in C leads to an increase in y (movement from points A to P) or an increase in C leads to a decrease in x (movement from points A to B). If y increases, then $yf_3(q_L)$ definitely increases and if x decreases, then $xf_4(I, q_L)$ definitely decreases. In either case, it follows from (25) that an increase in C causes $xf_4(I, q_L)$

constructing endogenous growth models is to show that public policy choices have long-run growth effects and semi-endogenous growth models effectively "throw the baby out with the bath water". However, given Steger's results about low speed of convergence, this criticism lacks practical relevance. In semi-endogenous growth models, public policies have growth effects that last a long time, they just don't last forever.

to fall. An increase in the rate of copying is associated with a decrease in the share of total employment in R&D activities. Summarizing, I have established

Theorem 2 *A permanent increase in the rate of copying C leads to a permanent decrease in the share of total employment in R&D activities and has a theoretically ambiguous effect on the rate of technological change.*

An increase in the rate of copying C directly cuts into the profits earned by innovating firms and contributes to more demand for production workers (since less monopoly pricing is associated with increased sales). If these were the only two considerations, then an increase in the rate of copying would definitely lead to a slowdown in the rate of technological change. However, an increase in the rate of copying also contributes to a more efficient utilization of R&D resources between leaders and followers (since firms earning monopoly profits have relatively weak incentives to innovate). Thus an increase in the rate of copying has a theoretically ambiguous effect on the rate of technological change.

Theorem 2 is useful for understanding the welfare results derived in section 3. An increase in the rate of copying leads to a decrease in the share of total employment in R&D activities and makes the case for subsidizing R&D stronger. Without copying ($C = 0$), the model tends to generate a R&D employment share that is unreasonably high and thus copying plays a central role in the welfare analysis.

2.10 Firm Value, Firm Size and R&D Effort

Focusing on a single steady-state equilibrium path, the model has interesting implications for how firm value, firm size and R&D effort vary across industries and over time.

Consider first the firm value v_L of single industry leaders (the firms in the model that earn monopoly profits). As an initial condition, I assume that $j(\omega, t) = 0$ for all industries ω at time $t = 0$. Then equation (14) implies that all of these industry leaders start off with the same market value at time $t = 0$ and more generally, market value takes the form $v_L = c\delta^{j(\omega,t)}$ where $c \equiv \frac{1-s}{\delta A_F}$ is a constant. Since the same innovation rate I prevails in all industries and is constant over time, the cross-sectional distribution of j is Poisson with parameter It. The mean of this distribution is It and the variance is also It, so both the mean and the variance of j increase over time. In the Appendix, I show that the distribution of v_L across industries becomes approximately lognormal when t is large. Given (15), the same holds for the distribution of v_{LF} across industries.

Next, consider how the firm size of single quality leaders varies across industries and over time. Measuring firm size by total revenue R, equation (5) implies that $R = \frac{p_L y}{x}\delta^{j(\omega,t)}$, where both x and y are constant over time. Using the same reasoning as above, because the cross-sectional distribution of j is Poisson with parameter It, the distribution of firm size R becomes approximately lognormal when t is large.

Finally, consider how R&D effort varies across industries. For single quality leaders, equation (8) implies that R&D expenditure is $(1-s)\ell_L = (1-s)\left(\frac{I_L}{A_L}\right)^{1/\beta}\delta^{j(\omega,t)}$. Since I_L is constant over time, R&D expenditure only varies across industries because of the $\delta^{j(\omega,t)}$ term. Using the same reasoning as above, I conclude that the distribution of R&D expenditure by single industry leaders becomes approximately lognormal when t is large. Furthermore, R&D intensity (defined by the ratio of R&D expenditure to total revenue) is constant across single industry leaders and over time. Thus the model implies that R&D intensity is independent of firm size. This is one of the stylized facts that have emerged from empirical studies using firm-level data [see Klette and Kortum (2004)] and is perhaps the central property of the Klette-Kortum model of innovating firms.

3 Welfare Analysis

In this section, the welfare implications of the model are explored, assuming that the social planner's objective is to maximize the discounted utility of the representative household. I assume that the social planner can only intervene by choosing the R&D subsidy rate $s(t)$ at each point in time t and solve for the welfare-maximizing steady-state R&D subsidy rate.

3.1 Utility Growth and Economic Growth

Taking into account that the price p_L is charged in m_L industries and the price 1 is charged in m_C industries, substituting (3) into (2) yields

$$u(t) = c(t)Q(t)^{\frac{1-\alpha}{\alpha}} f_5(q_L(t)) \tag{26}$$

where the f_5 function is defined by $f_5(q_L) \equiv [p_L^{1-\sigma} q_L + 1 - q_L]^{\frac{1-\alpha}{\alpha}}$. The static utility u earned by the representative consumer is increasing in consumer expenditure c, increasing in the average quality of products Q and decreasing in the degree of monopoly pricing q_L [f_5 is a decreasing function of q_L].

Given static utility, it is straightforward to calculate the steady-state rate of utility growth. Taking logs and differentiating (26) and then taking into account that both c and q_L are constants in a steady-state equilibrium yields

$$g \equiv \frac{\dot{u}(t)}{u(t)} = \frac{1-\alpha}{\alpha}\frac{\dot{Q}(t)}{Q(t)} = \frac{1-\alpha}{\alpha} n. \tag{27}$$

Like the steady-state innovation rate $I = \frac{n}{\delta - 1}$, the steady-state utility growth rate g is proportional to the population growth rate n.

Utility growth and economic growth are distinct concepts, but in this model, since the nominal wage equals one throughout time, and $u(t) = cQ(t)^{\frac{1-\alpha}{\alpha}} f_5(q_L)$ is the steady-state equilibrium utility at time t, $Q(t)^{\frac{1-\alpha}{\alpha}}$ is the real wage at time t. Thus, real wage growth coincides with utility growth along the steady-state equilibrium path and g is also the economic growth rate.

3.2 The Social Planner's Problem

Before solving for the welfare-maximizing R&D subsidy policy, I solve for the welfare-maximizing innovation rate policy. Using (1), (19) and (21), substituting for $y(t)$ using (23), substituting for $c(t)$ using (24), and substituting for $Q(t)$ using $x(t) \equiv \frac{Q(t)}{L(t)}$, the social planner's problem of maximizing the discounted utility of the representative household can be written as

$$\boxed{\begin{array}{l} \max_{I(\cdot)} \quad \int_0^\infty e^{-(\rho-n)t}\left\{\ln[1-xf_4(I,q_L)]+\ln f_5(q_L)-\ln f_6(q_L)+\frac{1-\alpha}{\alpha}\ln x\right\}\,dt \\ \qquad\qquad \text{subject to} \quad \dot{x}=f_1(x,I) \quad x(0)>0 \text{ given} \\ \qquad\qquad\qquad\qquad \dot{q}_L=f_2(I,q_L) \quad q_L(0)\in[0,1] \text{ given} \end{array}}$$

where I, q_L and x are all functions of time t and

$$f_6(q_L) \equiv \frac{p_L^{-\sigma} q_L + 1 - q_L}{p_L^{1-\sigma} q_L + 1 - q_L}.$$

This is a standard optimal control problem with two state variables x and q_L and one control variable I.[18]

I am interested in characterizing the initial level of relative R&D difficulty $x(0)$ such that there is nothing to gain from deviating from the steady-state innovation path [$I(t) = \frac{n}{\delta-1}$ for all t] when the initial value of q_L is its steady-state value $q_L(0) = \frac{I\delta}{I\delta+C}$. Once the welfare-maximizing steady-state value of x is determined, it is straightforward to infer the corresponding R&D subsidy rate that generates this x as a steady-state equilibrium outcome, since there is a monotonically increasing relationship between s and steady-state equilibrium x (Theorem 1). Since I_L, I_{LF}, and $\hat{\delta}$ are given by their steady-state equilibrium values in the definition of $f_4(\cdot)$, all possible ways of deviating are not allowed in the formulation of the optimal control problem. However, if it is not optimal to deviate from the steady-state path in any feasible way, then it is not optimal to deviate from the steady-state path when the allowable ways of deviating are restricted.

The current-valued Hamiltonian function for the social planner's optimal control problem is given by

$$\mathcal{H} \equiv \ln[1-xf_4(I,q_L)] + \ln f_5(q_L) - \ln f_6(q_L) + \frac{1-\alpha}{\alpha}\ln x + \theta_1 f_1(x,I) + \theta_2 f_2(I,q_L)$$

where θ_1 and θ_2 are co-state variables associated with the two state equations. Maximizing the current-valued Hamiltonian with respect to the control yields the first order condition

$$\frac{\partial \mathcal{H}}{\partial I} = \frac{-x/A_F}{1-xf_4(I,q_L)} + \theta_1 x(\delta-1) + \theta_2 \delta(1-q_L) = 0.$$

[18]The term $\ln L$ is left out of the social planner's problem since it plays no role in determining the optimal control.

In a welfare-maximizing steady-state where x, I and q_L are all constant over time, the first order condition implies that both θ_1 and θ_2 must be constant over time. Solving the first order condition for θ_1 yields

$$\theta_1 = \frac{1}{\delta - 1}\left\{\frac{1}{A_F(1 - xf_4^*)} - \frac{\theta_2\delta(1 - q_L)}{x}\right\} \tag{28}$$

where the notation f_4^* means the steady-state value of the f_4 function. Equation (28) states that θ_1 is a function of x and θ_2 only.

The first costate equation is $\dot{\theta}_1 = (\rho - n)\theta_1 - \frac{\partial \mathcal{H}}{\partial x}$ or spelling things out more fully

$$\dot{\theta}_1 = (\rho - n)\theta_1 + \frac{f_4(q_L, I)}{1 - xf_4(q_L, I)} - \frac{1 - \alpha}{\alpha}\frac{1}{x} - \theta_1[(\delta - 1)I - n].$$

Taking into account the steady-state values of I and q_L, using (28) to substitute for θ_1, and recognizing that θ_1 is constant over time in the welfare-maximizing steady-state, the first costate equation implies that

$$\left\{\frac{\rho - n}{\delta - 1}\frac{1}{A_F} + f_4^*\right\}\frac{x}{1 - xf_4^*} = \frac{1 - \alpha}{\alpha} + \left\{\frac{\rho - n}{\delta - 1}\delta(1 - q_L)\right\}\theta_2. \tag{29}$$

Equation (29) contains only two unknowns (x and θ_2) and describes an upward-sloping line in $\left(\frac{x}{1 - xf_4^*}, \theta_2\right)$ space when $C > 0$.

The second costate equation is $\dot{\theta}_2 = (\rho - n)\theta_2 - \frac{\partial \mathcal{H}}{\partial q_L}$ or spelling things out more fully

$$\dot{\theta}_2 = (\rho - n)\theta_2 + \frac{x\frac{\partial f_4(q_L, I)}{\partial q_L}}{1 - xf_4(q_L, I)} - \frac{f_5'(q_L)}{f_5(q_L)} + \frac{f_6'(q_L)}{f_6(q_L)} + \theta_2(\delta I + C).$$

Taking into account the steady-state values of I and q_L, and recognizing that θ_2 is constant over time in the welfare-maximizing steady-state, the second costate equation implies that

$$\left\{\frac{\partial f_4(q_L, I)}{\partial q_L}\right\}\frac{x}{1 - xf_4^*} + \{\rho - n + \delta I + C\}\theta_2 = \frac{f_5'(q_L)}{f_5(q_L)} - \frac{f_6'(q_L)}{f_6(q_L)}. \tag{30}$$

Equation (30) also contains two unknowns (x and θ_2) and describes an downward-sloping line in $\left(\frac{x}{1 - xf_4^*}, \theta_2\right)$ space since $\frac{\partial f_4(q_L, I)}{\partial q_L} > 0$.

Both the upward-sloping first costate equation (29) and the downward-sloping second costate equation (30) are illustrated in Figure 5. The intersection of these two lines at point A pins down the welfare-maximizing steady-state value of x, from which one can determine the welfare-maximizing steady-state R&D subsidy rate.

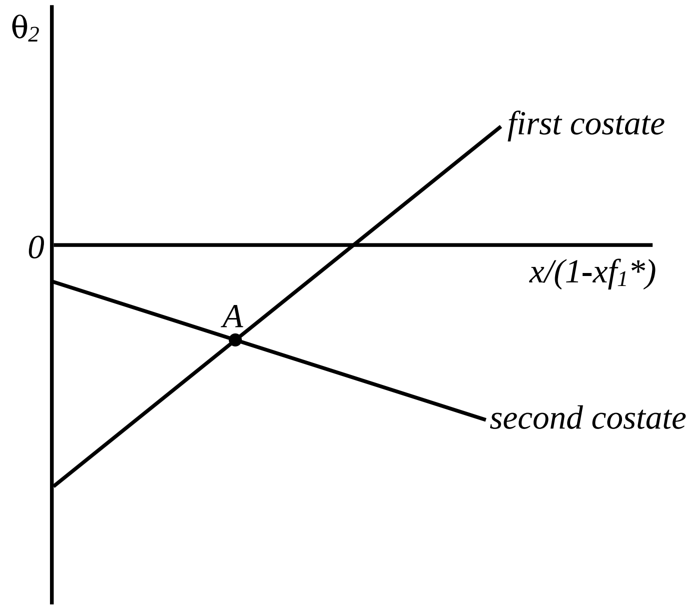

Figure 5: The welfare-maximizing steady-state

3.3 The No Copying Special Case

In general, equations (29) and (30) are too complicated to yield an analytical solution for the welfare-maximizing steady-state R&D subsidy rate but it is straighforward to solve these equations numerically using a computer. Before I do so, I want to focus on one special case where the model does yield a simple closed form solution for the welfare-maximizing steady-state R&D subsidy rate, namely, when the rate of copying C is zero.

When $C = 0$ and $q_L = 1$, then the first costate equation (29) is a vertical line in $\left(\frac{x}{1-xf_4^*}, \theta_2\right)$ space and the second costate equation (30) is not needed to determine the steady-state value of x. The first costate equation simplifies to

$$\left\{\frac{\rho - n}{\delta - 1}\frac{1}{A_F} + f_4^*\right\}\frac{x}{1 - xf_4^*} = \frac{1-\alpha}{\alpha}$$

which uniquely determines x. When $C = 0$, the resource condition (25) simplifies to $1 = y + xf_4^*$ which uniquely determines y given x. Substituting for y and x into the steady-state R&D condition (16) and simplifying using (12) and $p_L = \frac{1}{\alpha}$ yields

$$1 - s = \frac{\rho - \gamma I_L}{\eta(\rho + I) - \gamma I_L} > 0 \tag{31}$$

where $\eta \equiv \frac{\delta - 1}{\delta}$, $\gamma \equiv (\delta - 1)(1 - \beta\eta)$, I_L is given by (12) and $I = \frac{n}{\delta - 1}$. Equation (31) determines the welfare-maximizing steady-state R&D subsidy rate s. Since $\rho > n$ implies that $\rho > \eta(\rho + I)$ and $I > I_L$ implies that $\rho > \gamma I_L$, it follows from (31) that $1 - s > 1$ and $s < 0$. Summarizing, I have established

Theorem 3 *When there is no copying ($C = 0$), the welfare-maximizing steady-state public policy is a tax on R&D expenditures ($s < 0$).*

3.4 Numerical Results

When there is copying by firms of other firms' products, equations (29) and (30) are too complicated to solve analytically so I resort to computer simulations to gain some insight into the implications of the model. It is straighforward to solve these equations numerically for the welfare-maximizing steady-state x using a computer. Then the steady-state R&D condition (16) and the steady-state resource condition (25) can be solved simultaneously to determined the corresponding welfare-maximizing steady-state R&D subsidy rate s.

In the computer simulations, I use as benchmark parameter values $\rho = .04$, $n = .018$, $\alpha = .8$, $\lambda = 1.08$, $\beta = .3$, $A_F = 1$ and $A_L = .18$. The parameter choice $\rho = .04$ is a standard choice in the macro literature and means that the steady-state market interest rate is 4 percent. The population growth rate $n = .018$ (or 1.8 percent) corresponds to the world population growth rate during the 1980s [see Kremer (1993)]. The consumer demand parameter $\alpha = .8$ means that pure monopoly pricing in each industry involves a 25% markup of price over marginal cost and together with $n = .018$ implies that $g = \frac{1-\alpha}{\alpha} n = .0045$, that is, the steady-state economic growth rate is 0.45 percent. This economic growth rate is consistent with Denison (1985, p.30), who studied the US over the period 1929-82 and found the rate of GNP growth attributable to advances in knowledge to be of the order of one half percent per year. $\lambda = 1.08$ means that each innovation is a 8% improvement in product quality and implies that $\delta = \lambda^{\frac{\alpha}{1-\alpha}} = 1.36$, that is, each innovation increases R&D difficulty by 36%. The values of n and δ in turn imply that $I = \frac{n}{\delta - 1} = .05$ or $\frac{1}{I} = 20$, that is, an innovation occurs once every 20 years in the typical industry. For the decreasing returns to leader R&D parameter β, values between 0.1 and 0.6 have been suggested in the empirical literature, for instance by Kortum (1993). $\beta = .3$ represents an intermediate value in this range. The follower R&D parameter choice $A_F = 1$ is a convenient normalization.

The leader R&D parameter $A_L = .18$ is then chosen to guarantee that the assumption $I_{LF} < I$ is satisfied with a margin, that is, not all innovating is done by large firms. With these parameter choices, $I_L = .029$, $I_{LF} = .042$ and $I = .05$, that is, copied quality leaders are more innovative than single quality leaders and some R&D is always done by R&D followers (small firms).[19]

In the computer simulations, I consider three values for the rate of copying parameter: 0, .05 and .1. $C = 0$ represents no copying, $C = .05$ is an intermediate value and $C = .1$ is the main case of interest. When $C = .1$, the typical new product is copied after 10 years ($\frac{1}{C} = 10$), which is in line with the empirical evidence reported in Jones and Williams (2000). When $C = .05$, the typical new product is copied after 20 years ($\frac{1}{C} = 20$).[20] The results of the computer simulations are reported in Table 2. For each of the three values of C, Table 2 reports the steady-state equilibrium outcome when there is no government intervention ($s = 0$) and the steady-state equilibrium outcome when the government optimally intervenes (the welfare-maximizing R&D subsidy rate s). While the results in Table 2 should not be taken too seriously, these simulations illustrate the types of outcomes that the model is capable of generating.

Starting with the case of no copying ($C = 0$), the first row of numbers in Table 2 shows what happens when R&D is not subsidized ($s = 0$) and the second row of numbers shows what happens when R&D is optimally subsidized. The results are shocking: the welfare-maximizing steady-state R&D subsidy rate is a 115 percent tax ($s = -1.15$), that is to say, for every dollar that a firm spends on R&D, it should have to pay the government one dollar and fifteen cents in taxes. Table 2 also reveals a related problem with studying the no copying case: when there is no government intervention ($s = 0$), the steady-state share of

[19]Consistent with these innovation rates, Malerba and Orsenigo (1995) in their empirical study of 33 European industries found that trustified capitalism (innovation by industry leaders) is a more accurate characterization than competitive capitalism (innovation by follower firms) for two-thirds of the industries studied.

[20]Of course, an industry leader may be driven out of business by follower firm innovation before its product is copied. The average duration of monopoly profits for a firm that innovates and becomes an industry leader is $1/(C + I_F)$ years. Thus, innovative firms on average earn monopoly profits for 47.6 years when $C = 0$, for 14.1 years when $C = .05$ and for 8.3 years when $C = .1$.

employment in R&D is 12.8 percent. This is much too high to be realistic and increasing the R&D subsidy rate above zero would make the R&D employment share even higher (Theorem 1). When R&D is optimally taxed ($s = -1.15$), the R&D share drops to 6.4 percent and consistent with Theorem 1, the introduction of the 115% R&D tax leads to a temporary slowdown in the rate of technological change (relative R&D difficulty x drops from 5.55 to 2.77).

When the rate of copying is increased to every 20 years on average ($C = .05$), the steady-state equilibrium properties change significantly. Now the steady-state equilibrium without government intervention ($s = 0$) involves a 4.5 percent R&D employment share and monopoly pricing in the fraction .576 of the industries in the economy (q_L drops from 1 to .576). Furthermore, optimal government intervention now involves a positive R&D subsidy of 20.5 percent ($s = .205$). Consistent with Theorem 1, the introduction of the 20.5% R&D subsidy leads to a temporary acceleration in the rate of technological change (relative R&D difficulty x increases from 2.24 to 2.79) and a small increase in the R&D employment share (.045 to .056).

When the rate of copying is increased further to every 10 years on average ($C = .1$), which is the main case of interest, the steady-state equilibrium properties change significantly again. Now the steady-state equilibrium without government intervention ($s = 0$) involves a 2.7 percent R&D employment share and monopoly pricing in the fraction .404 of the industries in the economy (q_L drops from .576 to .404). With competitive pricing prevailing in 60% of the economy, the case for subsidizing R&D is much stronger. Optimal government intervention involves a R&D subsidy of 48.4 percent ($s = .484$). Consistent with Theorem 1, the introduction of the 48.4% R&D subsidy leads to a large temporary acceleration in the rate of technological change (relative R&D difficulty x increases from 1.46 to 2.76) and a large relative increase in the R&D employment share (.027 to .052).

I draw three conclusions from the computer simulation results reported in Table 2. First, the welfare implications of the model are very sensitive to what is assumed about the rate of copying. Varying the rate of copying from $C = 0$ to $C = .1$ causes the optimal government intervention to change from a 115 percent R&D tax to a 48 percent R&D subsidy! Second, when no copying is assumed, which is a commonplace assumption in the literature on optimal R&D subsidy policy [Stokey (1995) and Jones and Williams (2000) are representative examples], then the model predicts an unrealistically high R&D employment share (12.8 percent).[21] Third, in the benchmark case ($C = .1$), the model has much more plausible implications. With copying occuring every ten years on average, the model generates a R&D employment share of 2.7 percent without government intervention and an optimal R&D subsidy rate of 48.4 percent. Empirical studies tend to support heavily subsidizing R&D activities. For example, Jones and Williams (1998) estimate for the US economy that optimal R&D investment is two to four times larger than actual R&D investment.

If the assumption that firms face a cost of maintaining unused production facilities is dropped, then the model has another equilibrium where single quality leaders practice limit pricing ($p_L = \lambda$) instead of monopoly pricing ($p_L = 1/\alpha$). The model has been solved in such a way that it is easy to calculate this other type of equilibrium. One simply substitutes λ for p_L in equations (16), (25), (29) and (30). Using the same benchmark parameter values as were used to calculate Table 2, Table 3 shows how the model's properties change when single quality leaders practice limit pricing. It is still the case that the optimal R&D subsidy rate increases significantly as one increases the rate of copying. The main differerence is that, for each rate of copying, it is now optimal to subsidize R&D at a higher rate. For example, when the rate of copying is $C = .1$ (the benchmark case) , the optimal R&D subsidy rate of 48.4% under monopoly pricing becomes 72.5% under limit pricing.

3.5 The Externalities Associated With R&D

To understand the welfare findings reported in Table 2, it is helpful to think about the external effects associated with R&D investment. There are four external effects present in the model.

First, there is the *consumer surplus effect*. Every time a firm innovates, consumers benefit because they can buy a higher quality product at the same price that they used to pay for a lower quality substitute. Furthermore these consumer benefits last forever because

[21] I conjecture that this problem is more widespread than is commonly recognized. When numerical results are presented in the literature on optimal R&D subsidy policy, information about the R&D employment share is seldom reported.

future innovations build on all the innovations of the past. Because firms do not take into account in their profit-maximization calculations that consumers benefit from innovations, these consumer surplus gains represents one reason why firms may under-invest in R&D activities from a social perspective.

Second, there is the *across-industry business stealing effect*. Every time a firm innovates, it takes some demand away from quality leaders in other industries and lowers the profits earned by these firms. Because firms do not take into account in their profit-maximization calculations that innovations lower demand in other industries and consequently lower profits earned by other quality leaders, these reduced profits represent one reason why firms may over-invest in R&D activities from a social perspective.

Third, there is the *intertemporal R&D spillover effect*. Every time a firm innovates, innovating becomes more difficult. Future innovations become costlier to discover and more resources must go into the R&D activities to maintain the steady-state innovation rate. That means that less resources are left for producing consumer goods and consumer expenditure must decrease as a consequence. The decrease in consumer expenditure implies that the profits earned by industry leaders in other industries decrease. Because firms do not take into account in their profit-maximization calculations that each increase in R&D difficulty contributes to reducing the profits earned by other industry leaders, this represents a second reason why firms may over-invest in R&D activities from a social perspective.

Finally, there is the *within-industry business stealing effect*. Every time a follower firm innovates in an industry with a single quality leader, the previous quality leader experiences a windfall loss. By itself, this loss in profit income contributes to lower aggregate consumer income and expenditure, and hence lower profits for all industry leaders. Because followers do not take into account in their profit-maximization calculations that quality leaders sometimes experience windfall losses when followers innovate, these losses represent a third reason why firms may over-invest in R&D activities from a social perspective.

When there is no copying ($C = 0$), the one positive external effect associated with R&D (consumer surplus) is dominated by the three negative external effects (across-industry business stealing, intertemporal R&D spillover, within-industry business stealing), implying that it is optimal to tax R&D activities (Theorem 3). As the rate of copying is increased above zero, all of the four above-mentioned external effects continue to be present. However, because firms that innovate earn monopoly profits for a shorter expected period of time, both the across-industry and within-industry business stealing external effects become weaker when there is a faster rate of copying. This is the reason why the optimal R&D subsidy rate s increases as C increases in Table 2.

Jones and Williams (2000) assume that there is no copying in their calibrated endogenous growth model but nevertheless find that it is typically optimal to subsidize R&D. To understand what is driving their different results, it is helpful to first focus on the special case where there is no creative destruction due to innovation clusters, no intertemporal knowledge spillovers and no R&D duplication. In this special case, it is unambiguously optimal to subsidize R&D,[22] essentially because two of the above-mentioned negative external effects are not present: there is no within-industry business stealing effect (because innovations are new varieties) and no intertemporal R&D spillover effect.

One way that Jones and Williams (2000) depart from this special case is by assuming that each innovating firm earns profits not only by selling its new variety but also by selling some previously existing varieties (a form of creative destruction that I do not study). This change would appear to make overinvestment in R&D more likely but actually it has the opposite effect because all firms must charge lower markups to insure the adoption of their products when “innovation clusters” are larger.[23] A second way that Jones and Williams depart from the above-mentioned special case is by allowing for intertemporal knowledge spillovers. In their computer simulations, this change typically further expands the region of R&D underinvestment because the intertemporal knowledge spillovers are usually positive (knowledge created by researchers today contributes to the productivity of future researchers, implying a growing patents-per-researcher ratio over time). The third and remaining way that Jones and Williams depart from the above-mentioned special case is by allowing for R&D duplication (a negative external effect associated with R&D that I do not

[22]This can be seen by comparing equations (12) and (14) in Jones and Williams (2000) and then noting that the R&D share for the decentralized economy is less than the optimal R&D share when $\Psi = 0$, $\phi = 0$, and $\lambda = 1$.

[23]As Jones and Williams (2000) discuss on p. 79, increasing Ψ expands the region of underinvestment.

study). It is the presence of this R&D duplication externality that is the key to why Jones and Williams obtain a region of overinvestment in R&D. They find that it is only optimal to tax R&D when the negative R&D duplication external effect is sufficiently strong.

4 Conclusions

This paper presents an endogenous growth model that is designed to be roughly consistent with the experience of high-tech firms like Intel. In the model, industry leaders invest in R&D to improve their products and small firms invest in R&D to become industry leaders. Thus, model can account for Intel's experience: repeatedly innovating over the past 30 years to maintain its industry leadership position and developing the world's first microprocessor in 1971 when it was a small firm. Empirical evidence indicates that both small and large firms play important roles in producing innovations.

The model has a unique steady-state equilibrium where the innovation rate is constant over time and innovating becomes progressively more difficult over time. Starting from this steady-state equilibrium, a permanently higher R&D subsidy rate leads to a temporary increase in the innovation rate and permanently increases the share of total employment in R&D activities (Theorem 1). In contrast, a permanently higher copying rate permanently decreases the share of total employment in R&D activities (Theorem 2). These properties are quite intuitive. For example, firms immediately respond to the R&D subsidy rate increase by devoting more resources to R&D activities and the innovation rate jumps up. But with the complexity of the problems researchers are trying to solve growing over time at a faster than usual (steady-state) rate, relative R&D difficulty increases over time and this acts as a brake on the economy. The processes of gradually declining innovation rates and gradually increasing relative R&D difficulty continue until the economy reaches an outcome that is sustainable in the long run, with the original steady-state innovation rate but a larger R&D employment share.

Even though R&D subsidies do not have long-run growth effects, it does not follow that a laissez faire public policy is welfare maximizing. In fact, it is almost always optimal for the government to intervene. This paper solves for the optimal R&D subsidy rate:

what a social planner would choose to maximize the discounted utility of the representative household. In the special case where there is no copying, it is unambiguously optimal to tax R&D expenditures (Theorem 3) and for plausible parameter values, I obtain a R&D tax rate that is surprisingly high: 115 percent! However, when the rate of copying is increased, it eventually becomes optimal to subsidize R&D activities. When copying occurs every 10 years on average, which is in line with the empirical literature, the model generates an optimal R&D subsidy rate of 48 percent.

One drawback of the model is that the rate of copying is exogenous. In reality, firms devoted considerable resources to copying other firms' products. For example, Mansfield, Schwartz and Wagner (1981) found that imitation costs represent 65 percent of innovation costs in their study of 48 product innovations. An important topic for future research is to explore how the welfare implications change when copying is costly and public policy choices influence how much imitative R&D is done. The insights developed in this paper may prove useful in studying richer models with both costly innovation and costly imitation.

Appendix

The Solution to the Bellman Equations

Taking $v_F = v_{FL} = 0$ and equation (7) into account, the first order condition for an interior solution to the follower's Bellman equation (11) is $-(1-s) + A_F v_L(j_\omega + 1)/\delta^{j_\omega} = 0$ which, when solved for v_L yields (14). As was claimed earlier, $v_L(\cdot)$ only depends on j_ω and not separately on ω, t or other state variables. Furthermore, it is easily verified that the Bellman equation (11) is satisfied when (14) holds and $v_F = v_{FL} = 0$.

The first order condition for an interior solution to the single quality leader's Bellman equation (9) is $-(1-s) + [v_L(j_\omega + 1) - v_L(j_\omega)]\, \partial I_L / \partial \ell_L = 0$ which, when solved for I_L using (8) and (14) yields (12). Taking $v_F = v_{FL} = 0$ into account, the first order condition for an interior solution to the copied R&D leader's Bellman equation (10) is $-(1-s) + \frac{\partial I_{LF}}{\partial \ell_{LF}} [v_L(j_\omega + 1) - v_{LF}(j_\omega)] = 0$ which, when solved for I_{LF} using (8) and

(15) yields (13).

Substituting into the follower's Bellman equation (10) and simplifying yields (17). Using (13), it is straightforward to show that $f(\hat{\delta}) = A_L \left[\frac{A_L \beta}{A_F}\left(1 - \frac{1}{\hat{\delta}}\right)\right]^{\frac{\beta}{1-\beta}} \left[(1-\beta)\hat{\delta} + \beta\right]$, from which it follows that $f(\hat{\delta})$ is increasing in $\hat{\delta} \geq 1$, $f(1) = 0$ and $f(+\infty) = +\infty$. It will be shown later that in any steady-state equilibrium, $r = \rho$ and $I = n/(\delta - 1)$. Thus, (17) implies that $\hat{\delta} > 1$ is completely pinned down by parameter values and does not vary with s or C.

Substituting into the remaining single leader's Bellman equation (9) and simplifying using (5), (8), (14) and (15) yields (16). The only unknowns to be determined in (16) are x and y. It will be shown later that x and y are both constant over time in any steady-state equilibrium.

To determine the relationship between $\hat{\delta}$ and δ, suppose for the moment that $\hat{\delta} \leq \delta$. Then $f(\hat{\delta}) \leq f(\delta)$ since $f(\cdot)$ is a monotonically increasing function. Also taking into account that $\delta A_F \frac{p_L - 1}{1-s}\frac{y}{x} > 0$ and $C\left(\frac{\delta - \hat{\delta}}{\hat{\delta}}\right) \geq 0$, (17) implies that (16) cannot be satisfied. Thus, it must be the case that $\hat{\delta} > \delta$ in any steady-state equilibrium and it follows that $I_{LF} > I_L > 0$.

It only remains to be shown that $I > I_{LF}$ when $A_L > 0$ is sufficiently small. Equations (13) and (17) imply that a decrease in A_L raises $\hat{\delta}$ and reduces I_{LF}. Furthermore, as A_L converges to zero, $\hat{\delta}$ converges to plus infinity and I_{LF} converges to zero. Thus, given that $I = \frac{n}{\delta - 1}$ in any steady-state equilibrium, $I > I_{LF}$ holds if and only if $A_L > 0$ is sufficiently small.

The Distribution of Firm Value

In this subsection, I solve for the cross-sectional distribution of firm value v_L and how it evolves over time in a steady-state equilibrium.

A well known property of the Poisson distribution with parameter μ is that it converges to a normal distribution with mean μ and variance μ as μ converges to infinity. Thus, for sufficiently large t, the cross-sectional distribution of j becomes approximately normal with mean It and variance It. Now equation (14) implies that $v_L = c\delta^j$ where $c \equiv \frac{1-s}{\delta A_F}$

is a constant and $\ln v_L = \ln c + j \ln \delta$. Given that j is approximately normally distributed with mean It and variance It, $\ln v_L$ is approximately normally distributed with mean $\ln c + It \ln \delta$ and variance $(\ln \delta)^2 It$. Thus, the cross-sectional distribution of firm value v_L is approximately lognormal when t is large.

References

Aghion, P., Harris, C., Howitt, P. and Vickers, J. (2001), "Competition, Imitation and Growth with Step-by-Step Innovation," *Review of Economic Studies,* 68, 467-492.

Arrow, K. (1962), "Economic Welfare and the Allocation of Resources for Innovation," in Nelson, R., *The Rate and Direction of Innovative Activity.* Princeton: Princeton University Press.

Barro, R. and Sala-i-Martin, X. (1992), "Convergence," *Journal of Political Economy,* 100, 223-251.

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

Davidson, C. and Segerstrom, P. (1998), "R&D Subsidies and Economic Growth," *RAND Journal of Economics,* 29, 548-577.

Denison, E. (1985), *Trends in American Economic Growth, 1929-1982.* Washington, DC: The Brookings Institution.

Dinopoulos, E. and Thompson, P. (1999), "Scale Effects in Schumpeterian models of economic growth," *Journal of Evolutionary Economics,* 9, 157-185.

Etro, F. (2004), "Innovation by Leaders," *Economic Journal,* 114, 281-303.

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

Grossman, G. and Helpman, E. (1991b), *Innovation and Growth in the Global Economy* Cambridge, MA: The MIT Press.

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

Jones, C. (1995a), "Time series tests of endogenous growth models," *Quarterly Journal of Economics,* 110, 495-525.

Jones, C. (1995b), "R&D-Based Models of Economic Growth," *Journal of Political Economy,* 103, 759-784.

Jones, C. and Williams, J. (1998), "Measuring the Social Return to R&D," *Quarterly Journal of Economics,* 113, 1119-1135.

Jones, C. and Williams, J. (2000), "Too Much of a Good Thing? The Economics of Investment in R&D," *Journal of Economic Growth,* 5, 65-85.

Klette, T. and Kortum, S. (2004), "Innovating Firms and Aggregate Innovation," *Journal of Political Economy,* 112, 986-1018.

Kortum, S. (1993), "Equilibrium R&D and the Patent-R&D Ratio: U.S. Evidence," *American Economic Review, Papers and Proceedings,* 83, 450-457.

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

Kremer, M. (1993), "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics,* 108, 681-716.

Li, C. (2003), "Endogenous Growth Without Scale Effects: A Comment," *American Economic Review,* 93, 1009-1018.

Malerba, F. and Orsenigo, L. (1995), "Schumpeterian Patterns of Innovation," *Cambridge Journal of Economics,* 19, 47-65.

Malliaris, A. and Brock, W. (1982), *Stochastic Methods in Economics and Finance* Amsterdam: North Holland.

Malone, M. (1995), *The MicroProcessor: A Biography* New York: Springer-Verlag.

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

Romer, P. (1990), "Endogenous Technological Change," *Journal of Political Economy,* 98, S71-S102.

Scherer, F. (1980), *Industrial Market Structure and Economic Performance* 2nd ed., Chicago: Rand McNally.

Scherer, F. (1984), *Innovation and Growth: Schumpeterian Perspectives* Cambridge, MA: MIT Press.

Segerstrom, P. (1998), "Endogenous Growth Without Scale Effects," *American Economic Review,* 88, 1290-1310.

Segerstrom, P., Anant, T., and Dinopoulos, E. (1990), "A Schumpeterian Model of the Product Life Cycle," *American Economic Review,* 80, 1077-1092.

Segerstrom, P. and Zolnierek, J. (1999), "The R&D Incentives of Industry Leaders," *International Economic Review,* 40, 745-766.

Steger, T. (2003), "The Segerstrom Model: Stability, Speed of Convergence and Policy Implications," *Economics Bulletin,* 15, 1-8.

Stokey, N. (1995), "R&D and Economic Growth," *Review of Economic Studies,* 62, 469-489.

Thompson, P. and Waldo, D. (1994), "Growth and Trustified Capitalism," *Journal of Monetary Economics,* 34, 445-462.

Table 1: 2000 Net Sales and R&D Expenditures
of Select Industry Leaders (in billions of dollars)

Industry Leader	Net Sales	R&D Expenditure	R&D as a % of Sales
Boeing	$51.3	$1.4	2.8%
DuPont	$28.2	$1.7	6.3%
Eastman Kodak	$13.9	$0.7	5.6%
General Electric	$129.8	$2.1	1.7%
Hewlett Packard	$48.7	$2.6	5.4%
IBM	$88.4	$5.1	5.8%
Intel	$33.7	$3.9	11.5%
Johnson & Johnson	$29.1	$2.9	10.0%
Merck	$40.3	$2.3	5.8%
Microsoft	$22.9	$3.7	16.4%
Motorola	$37.5	$4.4	11.8%
3M	$16.7	$1.1	6.6%
Nokia	$27.0	$2.3	8.5%
Pfizer	$29.5	$4.4	15.0%
Xerox	$18.7	$1.0	5.6%

Table 2: Computer Simulation Results With Monopoly Pricing

Outcome	C	s	x	R&D share	q_L	Comments
Equilibrium	0	0	5.55	.128	1.00	No copying case
Optimal	0	-1.15	2.77	.064	1.00	
Equilibrium	.05	0	2.24	.045	.576	Intermediate case
Optimal	.05	.205	2.79	.056	.576	
Equilibrium	.1	0	1.46	.027	.404	Benchmark case
Optimal	.1	.484	2.76	.052	.404	

Table 3: Computer Simulation Results With Limit Pricing

Outcome	C	s	x	R&D share	q_L	Comments
Equilibrium	0	0	1.94	.045	1.00	No copying case
Optimal	0	.315	2.78	.064	1.00	
Equilibrium	.05	0	1.14	.023	.576	Intermediate case
Optimal	.05	.612	2.84	.057	.576	
Equilibrium	.1	0	0.82	.015	.404	Benchmark case
Optimal	.1	.725	2.88	.054	.404	