Journal of Economic Growth, 9, 81–123, 2004

Patents in a Model of Endogenous Growth

TED O'DONOGHUE

Department of Economics, Cornell University, 414 Uris Hall, Ithaca, NY 14853-7601, USA

JOSEF ZWEIMÜLLER

University of Zurich, IZA, CESifo, and CEPR

This paper examines patent protection in an endogenous-growth model. Our aim is twofold. First, we show how the patent policies discussed by the recent patent-design literature can influence R&D in the endogenous-growth framework, where the role of patents has been largely ignored. Second, we explore how the general-equilibrium framework contributes to the results of the patent-design literature. In a general-equilibrium model, both incentives to innovate and monopoly distortions depend on the proportion of industries that conduct R&D. Furthermore, patents affect the allocation of R&D resources across industries, and patents can distort resources away from industries where they are most productive.

Keywords: patent design, endogenous growth, quality improvement

JEL classification: L16, O31, O34

1. Introduction

The recent endogenous-growth literature—beginning with Romer (1990), Segerstrom et al. (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992)—has emphasized the role of industrial R&D for economic growth. In a general-equilibrium framework, growth is driven by technological progress achieved in private firms, and the literature asks what factors stimulate or retard growth. While the literature recognizes the need for patent protection to stimulate industrial R&D, there has been surprisingly little attention paid to the impact of patent policy on growth.[1]

The patent-design literature, on the other hand, addresses exactly the question of how does patent policy affect incentives for industrial R&D.[2] However, the patent-design literature has, for the most part, confined itself to partial-equilibrium analysis. This

1 For exceptions, see Segerstrom (1992), Davidson and Segerstrom (1993), Helpman (1993), Chou and Shy (1993), Cheng and Tao (1999), and Li (2001), although these papers examine protection against potential imitators and our focus shall be protection against future innovators.

2 There is a long line of research following Nordhaus (1969) examining patent design for isolated innovations. More recently, the literature has addressed patent design for cumulative innovation. For two-stage models, see Scotchmer (1991, 1996a,b), Green and Scotchmer (1995), Scotchmer and Green (1990), Chang (1995), Matutes et al. (1996), and Van Dijk (1996). For models of sequential innovation, see O'Donoghue et al. (1998), O'Donoghue (1998), and Hunt (1999).

deficiency seems particularly important: Since a single patent policy applies to multiple industries, an analysis of how policy affects a single industry seems incomplete.

In this paper, we attempt to merge these two literatures. The aim of the paper is twofold. First, we examine the role of patent policy in the context of endogenous growth. Second, we explore how the general-equilibrium framework contributes to the results of the patent-design literature.

The patent-design literature examines (at least) four tools of patent policy. Patent life is the length of time for which a patent is valid. A patentability requirement is a minimum innovation size required to receive a patent. A patent's breadth puts restrictions on the products other firms can produce without a license. Lagging breadth limits imitation by specifying inferior products that other firms cannot produce. Leading breadth limits future innovators by specifying superior products that other firms cannot produce. Notice that lagging breadth represents protection against imitation, whereas leading breadth and a patentability requirement represent protection against future innovators. See O'Donoghue (1998) for a more detailed discussion of how these policy instruments have been used in the literature as well as how these policy instruments relate to the existing patent law.

The endogenous-growth literature has recognized the need for patent protection to stimulate growth. For the most part, however, attention has been limited to one simple policy—often referred to as ''infinitely-lived patents''. More accurately, we interpret the endogenous-growth patent policy as infinitely-lived patents that prevent all imitation (i.e., there is complete lagging breadth), but allow any superior product to displace the innovator (i.e., there is no leading breadth and no patentability requirement).

In Section 2, we outline a simple model of endogenous growth along the lines of Grossman and Helpman (1991) and Aghion and Howitt (1992). Within this framework, we embed a model of patent policy, which includes the standard endogenous-growth patent policy but also allows for policies with protection against future innovators. In Section 3, we address our first question: What is the role of patents for endogenous growth? We take the endogenous-growth policy as a benchmark, and then show how protection against future innovators can stimulate R&D investment. Specifically, if there is underinvestment under the benchmark policy, then policy can stimulate R&D with either a patentability requirement (as suggested by O'Donoghue, 1998 and Hunt, 1999), or leading breadth (as suggested by O'Donoghue, 1998).

While our main analysis focusses for simplicity on a first-generation growth model, Jones (1995) rejects such models because they exhibit counterfactual ''scale effects''. In Section 4, we address whether our main conclusions would survive in the more elaborate second- and third-generation growth models that do not exhibit scale effects. In short, the answer is that, in principle, our policies can be effective in such models. However, they are not effective in many existing models due to the fact that these models eliminate the combination of R&D within product lines and leapfrogging on which our results depend. To the extent that this combination was eliminated for tractability, and not because it is viewed as unrealistic, our main conclusions are relevant for the endogenous-growth literature. Indeed, in Appendix C we present a third-generation growth model (based on Howitt, 1999 and Aghion and Howitt, 1998, chapter 12) in which our main conclusions hold.

In Section 5, we address our second question: What is the patent-design literature missing by using partial-equilibrium analysis? A partial-equilibrium analysis of patent policy ignores the fact that policy changes affect multiple industries. A general-equilibrium analysis enables us to incorporate this feature of patent policy. In doing so, we find two important factors missing from partial-equilibrium analyses. First, there is a general-equilibrium effect from policy changes. In a partial-equilibrium framework, stronger patents imply increased profits for successful firms. In a general-equilibrium framework, however, these increased profits imply a reduction of the real wage and hence in the real cost of R&D, which reinforces the effectiveness of patents. Moreover, because this effect is stronger the larger is the innovative sector, patent policies are more effective as more industries innovate. The second factor missing from most partial-equilibrium analyses is a formal model of the static inefficiency (or output distortions) associated with patents. The fact that multiple industries use patents can imply that the output distortions created by patents are small—indeed there may be essentially no output distortions in the extreme cases where almost all industries use patents or very few industries use patents. Hence, partial-equilibrium analyses may be overemphasizing the importance of monopoly distortions created by patents.

In Section 6, we extend our model to explore an issue largely ignored by both the endogenous-growth literature and the patent-design literature: What are the implications of there being asymmetric R&D capabilities across industries? The empirical R&D literature suggests there are significant cross-industry differences in R&D productivity (see the survey by Cohen and Levin, 1989). If there are asymmetric R&D capabilities across industries, then in addition to the aggregate level of R&D, policy must also concern itself with the allocation of R&D resources across industries. We find that the private equilibrium tends to distort R&D resources away from those industries where these resources are more productive. Furthermore, stronger patent protection can exacerbate these distortions. These results are driven by a higher rate of creative destruction in the more productive industries, which induces firms to invest less than desired in those industries.

We conclude in Section 7 by discussing some limitations of our model and also some general lessons to take away from our analysis.

2. A Model of Endogenous Growth

In this section, we lay out a model of endogenous growth that is similar to those in Grossman and Helpman (1991) and Aghion and Howitt (1992). Within this model we embed a model of patent policy that includes the standard endogenous-growth patent policy as a special case but also allows for protection against future innovators. We consider a simple economy where there are two types of industries. First, there is a high-technology sector—a set of industries that conduct R&D to improve product quality. Second, there is a noninnovative sector—a set of industries where quality improvements are not possible. For simplicity we assume that labor is the only productive input in all sectors and that there is a fixed set of (labor) resources. Within this setting there are two allocative questions: (i) how to allocate labor between production (for consumption) and R&D; and (ii) how to allocate production labor between the high-technology sector and the noninnovative sector. In addition, there is a third question that will affect the performance of this economy: (iii) How ambitious should the R&D projects be in the high-technology sector (i.e., should firms pursue small or large quality improvements)?

2.1. The Underlying Model

The underlying model has three components: the R&D process, intertemporal preferences, and the resource constraint.

Economic growth is driven by endogenous product improvements. There is a continuum of goods indexed by $\omega \in [0, 1]$, each produced within its own industry. Individual goods may be available in multiple qualities. For good ω, let $q_\omega(t)$ be the maximum technologically feasible quality at time t. At time t, all firms are capable of producing any quality $q \leq q_\omega(t)$ (i.e., imitation is costless), and no firm is capable of producing any quality $q > q_\omega(t)$.[3]

The evolution of $q_\omega(t)$ for each ω is determined by R&D behavior. Quality improvements occur in only a fraction of industries. Specifically, there is an $\bar{\omega} \in [0, 1]$ such that in industries with $\omega \in (\bar{\omega}, 1]$, $q_\omega(t) = 1$ for all t. We refer to these industries as the noninnovative sector. The industries with $\omega \in [0, \bar{\omega}]$ constitute the high-technology sector.[4] In each high-technology industry, firms conduct R&D to repeatedly increase the maximum feasible quality in that industry. For simplicity, we assume $q_\omega(0) = 1$ for each $\omega \in [0, \bar{\omega}]$, and that the i-th innovation in industry ω increases the maximum feasible quality by factor $\gamma_{i\omega} > 1$. Hence, if there have been exactly I innovations in industry ω before date t, then $q_\omega(t) = \Pi_{i=1}^{I} \gamma_{i\omega}$. The innovation size $\gamma_{i\omega}$ is endogenous, as described below.

Innovations occur according to a Poisson process. If a firm has arrival rate of innovations ϕ, then the date of success τ has cumulative distribution $F(\tau) = 1 - e^{-\phi\tau}$. Each firm's arrival rate depends on the number of research workers it hires and on the innovation size it pursues. If a firm hires n research workers and pursues innovation size γ, it will have Poisson arrival rate $\phi = \lambda(\gamma)n$, where $d\lambda/d\gamma < 0$ and $d^2\lambda/d\gamma^2 \leq 0$. There are constant returns to scale for R&D labor, and the Poisson arrival rate is decreasing in the innovation size—that is, larger innovations are more difficult to achieve. We assume that the research technology is identical in all industries and at all times. Furthermore, we assume that firms' R&D processes are independent, so the arrival rate of innovations in industry ω, denoted ϕ_ω, is the sum of the individual firms' arrival rates.

Product quality matters because consumers prefer to consume higher-quality goods. There are L consumers with identical intertemporal preferences that can be represented by the intertemporal utility function

---

3 We consider a model along the lines of Grossman and Helpman (1991) where product improvements occur in consumption goods. Everything is essentially the same in a model where product improvements occur in intermediate goods, as in Aghion and Howitt (1992).

4 In their basic model, Grossman and Helpman (1991) consider the case where all industries are innovative, or $\bar{\omega} = 1$.

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

where t is an index of continuous time, ρ is the rate of time preference, and $\ln u(t)$ is instantaneous utility at time t. The instantaneous utility function is

$$\ln u(t) = \int_0^1 \ln[q_\omega(t)\, x_\omega(t)]\, d\omega, \tag{2}$$

where $x_\omega(t)$ is the quantity consumed of quality $q_\omega(t)$ of good ω at date t. This formulation assumes that a person consumes only the maximum feasible quality from each industry. Even so, we require that the maximum feasible quality $q_\omega(t)$ have the lowest quality-adjusted price in industry ω (because otherwise consumers would purchase an inferior quality). In other words, if good ω is also available in quality q' at price p', then the price of quality $q_\omega(t)$, denoted $p_\omega(t)$, must satisfy $p_\omega(t)/q_\omega(t) \leq p'/q'$.[5]

The final component of the underlying model is the resource constraint. Each person supplies one unit of labor, so the total labor supply is L. There are constant returns to scale in output production, and in each industry labor is the only input with a unit labor requirement a for all qualities.[6] If each consumer consumes $x_\omega(t)$ from industry ω, total consumption from industry ω is $Lx_\omega(t)$, and the total labor requirement is $Lax_\omega(t)$. Economy-wide employment in production is therefore $\int_0^1 Lax_\omega(t)\, d\omega$. Let $N(t)$ be the total number of research workers economy-wide, and $n_\omega(t)$ be the number of research workers in industry ω, which implies $N(t) = \int_0^{\bar{\omega}} n_\omega(t)\, d\omega$. Then for all t the resource constraint is

$$L = La \int_0^1 x_\omega(t)\, d\omega + \int_0^{\bar{\omega}} n_\omega(t) d\omega = La \int_0^1 x_\omega(t)\, d\omega + N(t). \tag{3}$$

We can summarize the underlying model as follows. The exogenous parameters are the number of consumers L, the intertemporal time preference ρ, the labor requirement for production a, and the fraction of industries that are innovative $\bar{\omega}$. The endogenous variables are the labor allocations $x_\omega(t)$ and $n_\omega(t)$, and the innovation sizes $\gamma_{i\omega}$.

2.2. *The Social Optimum*

A social planner will choose the endogenous variables to maximize intertemporal utility (equations (1) and (2)) subject to the resource constraint (equation (3)). Given the stationary nature of the model, the socially optimal labor allocations will be stationary—

5 The instantaneous utility function and the assumption about when consumers purchase the maximum feasible quality represent a reduced form of underlying instantaneous utility function

$$\ln u(t) = \int_0^1 \ln\left[\int_0^{q_\omega(t)} q\, x_{\omega q}(t)\, dq\right] d\omega$$

where $x_{\omega q}(t)$ is the quantity consumed of quality q of good ω at date t.

6 We could allow for a different labor requirement in the noninnovative sector than in the high-technology sector. Such an assumption would not change the qualitative nature of our results.

that is, for all ω we have $x_\omega(t) = x_\omega$ and $n_\omega(t) = n_\omega$ for all t. Furthermore, the socially optimal innovation size will be constant, or $\gamma_{i\omega} = \gamma$ for all i and ω.

Equation (2) can be written as $\ln\, u(t) = \int_0^1 \ln\, x_\omega\, d\omega + \int_0^1 \ln\, q_\omega(t)\, d\omega$, and we can then write intertemporal utility as

$$U = \frac{\int_0^1 \ln\, x_\omega\, d\omega}{\rho} + E\int_0^1 \left[\int_0^\infty \mathrm{e}^{-\rho t} \ln\, q_\omega(t)\, dt\right] d\omega.$$

Since for any noninnovative industry $\omega \in (\bar{\omega}, 1]$, $q_\omega(t) = 1$ for all t, $\int_0^\infty \mathrm{e}^{-\rho t} \ln\, q_\omega(t)\, dt = 0$. For any high-technology industry $\omega \in [0, \bar{\omega}]$, if the stationary Poisson arrival rate of innovations is $\lambda(\gamma)n_\omega$, then we have[7]

$$\begin{aligned} E\int_0^\infty \mathrm{e}^{-\rho t} \ln\, q_\omega(t)\, dt &= \frac{\ln q_\omega(0)}{\rho + \lambda(\gamma)n_\omega} + \left(\frac{\lambda(\gamma)n_\omega}{\rho + \lambda(\gamma)n_\omega}\right) \frac{\ln[\gamma q_\omega(0)]}{\rho + \lambda(\gamma)n_\omega} \\ &\quad + \left(\frac{\lambda(\gamma)n_\omega}{\rho + \lambda(\gamma)n_\omega}\right)^2 \frac{\ln[\gamma^2 q_\omega(0)]}{\rho + \lambda(\gamma)n_\omega} + \cdots \\ &= \frac{\ln q_\omega(0)}{\rho} + \frac{\ln \gamma}{\rho}\frac{\lambda(\gamma)n_\omega}{\rho}. \end{aligned}$$

Since for each $\omega \in [0, \bar{\omega}]$ $q_\omega(0) = 1$, we have

$$\rho U = \int_0^1 \ln\, x_\omega\, d\omega + \frac{\ln \gamma}{\rho}\lambda(\gamma) \int_0^{\bar{\omega}} n_\omega\, d\omega = \int_0^1 \ln\, x_\omega\, d\omega + \frac{\ln \gamma}{\rho}\lambda(\gamma)N.$$

If the social planner allocates labor aX to production, each consumer receives total consumption $\int_0^1 x_\omega\, d\omega = X/L$. Given this constraint, and the identical production technology in all industries, $\int_0^1 \ln x_\omega d\omega$ is maximized by consuming equal quantities from all industries. Hence, for any X, a social planner will choose $x_\omega = X/L$ for all ω. The resource constraint then implies $X = (L - N)/a$, and the social planner's problem becomes choosing γ and N to maximize

$$\rho U = \ln\left(\frac{L - N}{La}\right) + \frac{\ln\, \gamma}{\rho}\lambda(\gamma)N.$$

If we let γ^* denote the socially optimal innovation size, and N^* denote the socially optimal level of aggregate R&D, then γ^* and N^* satisfy

7 This equation uses the following calculations:
(a) If the flow profit π is received until uncertain time t that has Poisson arrival rate ϕ, then it has expected value

$$\int_0^\infty \left(\pi \frac{1 - \mathrm{e}^{-\rho t}}{\rho}\right)\phi \mathrm{e}^{-\phi t}\, dt = \frac{\pi}{\rho + \phi}.$$

(b) If the payoff v is received at uncertain time t that has Poisson arrival rate ϕ, then it has expected value

$$\int_0^\infty (v\mathrm{e}^{-\rho t})\phi \mathrm{e}^{-\phi t}\, dt = \frac{\phi}{\rho + \phi} v.$$

$$\frac{\lambda(\gamma^*)}{-\,d\lambda/d\gamma} = \gamma^* \ln \gamma^* \tag{4}$$

and

$$N^* = L - \frac{1}{\ln \gamma^*} \frac{\rho}{\lambda(\gamma^*)}. \tag{5}$$

γ^* and N^* are both independent of $\bar{\omega}$, the fraction of industries that conduct R&D. Furthermore, the social planner cares only about aggregate R&D N^* and not how R&D labor is allocated across industries (i.e., any set of n_ω such that $\int_0^{\bar{\omega}} n_\omega \, d\omega = N^*$ will do). These results follow from the separability of quality and quantity in the instantaneous utility function, and the constant returns to scale in R&D. Separability implies consumers are indifferent to where quality improvements occur, and constant returns imply R&D is equally good in all industries.

2.3. Patent Policy

We analyze the performance of this economy relative to the social optimum when policymakers are constrained to create incentives for R&D with patents. Following O'Donoghue et al. (1998) and O'Donoghue (1998), when innovation improves products along a unidimensional quality measure, we can identify four distinct tools of patent design: Patent life, a patentability requirement, lagging breadth, and leading breadth. Patent life is the length of time for which a patent is valid. A patentability requirement specifies a minimum innovation size required to patent a new product. A patent's breadth specifies products that would infringe upon the patent, which means that these products cannot be produced without the patentholder's permission (in the form of a license). Lagging breadth specifies inferior products that infringe, and leading breadth specifies superior products that infringe. The patentability requirement and patent breadth reflect the two main tasks that confront the patent authorities. First, they must assess whether a new technology merits a patent—whether the invention satisfies the statutory requirements of novelty, nonobviousness, and utility. Second, they must decide which alternative technologies infringe on the patent—which means that the patentholder has the right to exclude other firms from using those technologies (and can sue them for damages if they do).[8]

Our main focus is whether protection against future innovators can stimulate R&D relative to patents without such protection. Except for occasional discussions, we simplify the analysis by fixing patent life and lagging breadth, so that the policy tools of interest are a patentability requirement and leading breadth. Specifically, we focus on patent policies

8 By ''patent authorities'', we mean the combination of the patent office and the patent courts. The patent office makes initial decisions on both dimensions, but these decisions are often modified by the patent courts during infringement suits. The courts sometimes rule that a patent should not have been granted—indeed, this is a common defense in infringement suits. In addition, the courts sometimes contract and sometimes expand the set of alternative technologies that are covered by a patent.

where patent life is infinite and there is ''complete lagging breadth''. In our model, patents are effectively terminated by future innovations, not by when the patent expires (O'Donoghue et al., 1998 define this as ''effective patent life''). Hence, infinite patent life is meant to be a proxy representing that there is a very small probability of a patent having value when it expires. ''Complete lagging breadth'' means that all qualities made feasible by an innovator are protected. In other words, if a firm has a patent on quality q_o and its innovation size was γ, then no other firm can produce any quality $q \in (q_o/\gamma, q_o]$ during the life of the patent without a license. An implication of infinite patent life and complete lagging breadth is that at all times the most recent innovator will produce the maximum feasible quality $q_\omega(t)$, and the nearest rival will be the previous innovator in that industry.

We denote a patentability requirement by $P \in [1, \infty)$, where a firm can receive a patent only if it has innovation size $\gamma \geq P$. A patentability requirement represents a lower bound on innovation size (as long as firms are interested in receiving a patent—see our discussion of trade secrets in Section 3). We denote leading breadth by $K \in [1, \infty)$ such that the patent on quality q_o prevents other firms from producing any quality $q \in [q_o, Kq_o)$ during the life of the patent without a license from the patentholder. We reiterate the distinction between a patentability requirement and leading breadth. Both put restrictions on future innovators, but a patentability requirement restricts what future innovators can patent, whereas leading breadth restricts what they can produce without infringing. Hence, for example, even if an innovator can get a patent on his new product, he may have to pay licensing fees to some previous innovators in order to produce (as we discuss in more detail in Section 3).

We denote a specific patent policy by $\psi \equiv (P, K)$. Under policy ψ, all patents have infinite patent life, complete lagging breadth, leading breadth K, and the patentability requirement for each generation is P. Different patent policies will induce different outcomes.

The endogenous-growth models of Grossman and Helpman (1991) and Aghion and Howitt (1992) assume a very simple patent policy: Each successful firm receives an infinitely-lived patent that prevents other firms from producing its quality. There is no explicit patent breadth, but these models effectively assume complete lagging breadth since product quality is discrete—discrete product quality implies that the nearest feasible competing product is the previous state-of-the-art product. Importantly, however, there is no patentability requirement and no leading breadth (i.e., there is no protection against future innovators). Hence, we interpret the endogenous-growth patent policy as the policy $(P = 1, K = 1) \equiv \psi_o$.[9]

9 A few papers consider weaker patent policy. For example, Segerstrom (1992), Davidson and Segerstrom (1993), and more recently Aghion et al. (2001) explore patent protection where they assume stronger patent protection implies a decreased probability of imitation, and Helpman (1993) explores property rights where he assumes that tighter property rights imply imitation is more costly. With our terminology, these papers explore lagging breadth, and assume no leading breadth.

2.4. The Private Equilibrium under Policy ψ_o

We now examine how the endogenous variables—$x_\omega(t)$ and $n_\omega(t)$ for each ω and t, and $\gamma_{i\omega}$ for each i and ω—are determined in private markets under patent policy ψ_o. We will use this outcome as a benchmark against which to compare policies with protection against future innovators in Section 3. To solve the model, we examine output markets (from which all profits are derived) and R&D markets (where firms innovate in order to gain an advantageous output-market position).

We begin with the demand functions for consumption goods. Consumers maximize utility (equations (1) and (2)) subject to their budget constraint. We suppose all consumers have identical wealth $A(0)$ (e.g., all consumers own equal shares of all firms). In addition, each consumer supplies one unit of labor at all times that earns wage w. Throughout, we take the wage w to be the numeraire, although for expositional clarity we often include w in equations. At time t, for each industry ω each consumer purchases quantity $x_\omega(t)$ of quality $q_\omega(t)$ at price $p_\omega(t)$. Letting r denote the rate of interest, each consumer's budget constraint is

$$\int_0^\infty \mathrm{e}^{-rt}\left[\int_0^1 p_\omega(t)\,x_\omega(t)\,d\omega\right]dt \le A(0) + \int_0^\infty \mathrm{e}^{-rt} w\,dt.$$

Wealth evolves according to $\dot{A}(t) = [rA(t) + w] - \left[\int_0^1 p_\omega(t)\,x_\omega(t)\,d\omega\right]$. Restricting attention to balanced growth steady states where $\dot{A}(t) = 0$ for all t, we have $r = \rho$, and for all t instantaneous income is $w + \rho A(0) \equiv Y$. Each consumer spends exactly Y at any date t, and chooses consumption bundle $\{x_\omega(t)\}_{\omega\in[0,1]}$ to maximize instantaneous utility. Given symmetric Cobb–Douglas utility, consumers allocate income Y to each industry ω, which means that for each industry ω each consumer demands quantity $Y/p_\omega(t)$. Since there are L consumers, the demand function for each consumption industry is $LY/p_\omega(t)$.

Next, we examine the optimal behavior of firms in output markets. We assume each noninnovative industry $\omega \in (\bar{\omega}, 1]$ is competitive. Firms price at marginal cost, so given labor requirement a and wage w, $p_\omega(t) = wa$ for all t. Hence, at all times each consumer purchases $x_\omega(t) = Y/(wa)$ for each $\omega \in (\bar{\omega}, 1]$.

In each high-technology industry $\omega \in [0, \bar{\omega}]$, patent protection creates imperfect competition. At time t, the most recent innovator produces quality $q_\omega(t)$, and patent protection will determine the largest quality $q' < q_\omega(t)$ that a rival can produce without a license. We shall refer to the most recent innovator as the market leader. The market leader will serve the entire market, but the price $p_\omega(t)$ is constrained in that it must be the lowest quality-adjusted price in industry ω. Since rivals are willing to price at marginal cost wa, the market leader's price will be $p_\omega(t) = \mu wa$ where the markup $\mu = q_\omega(t)/q'$. As we shall see, the markup μ will be the same in all high-technology industries and at all times. Given demand function $LY/p_\omega(t)$, at all times the market leader in each industry $\omega \in [0, \bar{\omega}]$ earns profit $\pi = LY((\mu - 1)/\mu)$.

We next examine behavior in R&D markets. We assume perfect competition in the market for R&D. As a result, it will turn out that the market leader in industry ω will not conduct R&D in industry ω. As in Grossman and Helpman (1991) and Aghion and Howitt (1992), it is more profitable to gain a one-step advantage than to extend a one-step advantage to a two-step advantage. Consider the payoff to an R&D firm that chooses innovation size γ and hires R&D labor n. The firm's instantaneous wage cost is wn. The firm has success with Poisson arrival rate $\lambda(\gamma)n$. If we let V represent the reward to success, then the firm's expected instantaneous payoff is $-wn + \lambda(\gamma)nV$.

What is the reward to success V? When market leaders do not conduct R&D, at all times every market leader has a one-step advantage. This means that if a firm has an innovation of size γ, the firm is able to charge markup $\mu = \gamma$ (given complete lagging breadth). Hence, an R&D firm will earn a profit flow $\pi = LY((\gamma - 1)/\gamma)$ following a success. Since market leaders do not conduct R&D, the reward to success consists of earning flow profit π until the first subsequent success in the industry. If ϕ is the equilibrium Poisson arrival rate in the industry, then this flow of profit has discounted expected value $\pi/(\rho + \phi)$ (see the calculation in footnote 7). In other words, the reward to success $V = \pi/(\rho + \phi)$.

We can now rewrite an R&D firm's instantaneous payoff as

$$-wn + \lambda(\gamma)n\frac{\pi}{\rho + \phi} = -wn + \lambda(\gamma)nLY\left(\frac{\gamma - 1}{\gamma}\right)\frac{1}{\rho + \phi},$$

where the R&D firm takes L, Y, w, ρ, and ϕ as given and chooses γ and n. Since all R&D firms in all industries will choose γ to maximize $\lambda(\gamma)((\gamma - 1)/\gamma)$, at all times all firms conducting R&D will choose innovation size γ_o defined by

$$\frac{\lambda(\gamma_o)}{-d\lambda/d\gamma} = \gamma_o(\gamma_o - 1). \tag{6}$$

Since there are constant returns to scale in the number of research workers n, the individual research venture is of indeterminate size. Even so, free entry requires that if $n > 0$ then

$$\lambda(\gamma_o)\frac{\pi}{\rho + \phi} = w. \tag{7}$$

We shall denote the private equilibrium level of R&D labor by $N_o = \bar{\omega}n_o$, where n_o is the number of R&D workers hired in each high-technology industry and N_o is the number of R&D workers hired economy-wide.

At this point, it is convenient to combine the demand functions for consumption goods, optimal firm behavior in output markets, and the resource constraint (equation (3)) into a single equation that expresses market profits π as a function of the markup μ, aggregate R&D N, and exogenous parameters. As described above, for all t, $x_\omega(t) = Y/(\mu wa)$ for $\omega \in [0, \bar{\omega}]$ and $x_\omega(t) = Y/(wa)$ for $\omega \in (\bar{\omega}, 1]$. Plugging these equations into equation (3), taking the wage w to be the numeraire, and solving for instantaneous income Y, we obtain $Y = ((L - N)/L)(\mu/((1 - \bar{\omega})\mu + \bar{\omega}))$. Since $\pi = LY((\mu - 1)/\mu)$, we can conclude that

$$\pi = \frac{(L - N)(\mu - 1)}{(1 - \bar{\omega})\mu + \bar{\omega}}. \tag{8}$$

Note that equation (8) is a condition that must hold in equilibrium, but does not represent market profits as perceived by individual firms. Substituting π from equation (8) into the no-profit condition (equation (7)), we derive the following expression for N_o in terms of exogenous parameters:

$$N_o = \bar{\omega} n_o = \begin{cases} \bar{\omega}\left[\frac{\gamma_o - 1}{\gamma_o} L - \frac{(1-\bar{\omega})\gamma_o + \bar{\omega}}{\gamma_o} \frac{\rho}{\lambda(\gamma_o)}\right] & \text{if } L \geq \frac{(1-\bar{\omega})\gamma_o + \bar{\omega}}{\gamma_o - 1} \frac{\rho}{\lambda(\gamma_o)} \\ 0 & \text{if } L \leq \frac{(1-\bar{\omega})\gamma_o + \bar{\omega}}{\gamma_o - 1} \frac{\rho}{\lambda(\gamma_o)}. \end{cases} \tag{9}$$

We summarize this outcome in Lemma 1.

Lemma 1 Under policy ψ_o:

i. The market leader in industry $\omega \in [0, \bar{\omega}]$ does not conduct R&D.

ii. Firms choose innovation size γ_o defined by equation (6).

iii. The markup in industry $\omega \in [0, \bar{\omega}]$ is $\mu = \gamma_o$.

iv. The equilibrium level of R&D N_o is given by equation (9).

Consider how the outcome under policy ψ_o compares to the socially optimal outcome. A comparison of equation (6) to equation (4) reveals that $\gamma_o < \gamma^*$. And a comparison of equations (5) and (9) reveals that for any $\bar{\omega}$ the private market outcome can involve too little or too much R&D. These results are identical to those in Grossman and Helpman (1991) and Aghion and Howitt (1992), and the intuition is as described there.

2.5. *Constrained Social Welfare*

Patents create markups in the high-technology sector, whereas price equals marginal cost in the noninnovative sector. These markups distort consumption towards the noninnovative sector. We close this section by deriving a constrained social welfare function that takes these markups into account. Suppose a patent policy induces economy-wide R&D N, but creates markup μ in each high-technology industry. Then consumption from each noninnovative industry is $x_N \equiv Y/(wa)$, and consumption from each high-technology industry is $x_H \equiv Y/(\mu wa)$. Plugging $x_N = \mu x_H$ into the resource constraint (equation (3)) yields $x_H = (1/((1-\bar{\omega})\mu + \bar{\omega}))(L-N)/(La)$. We can then write intertemporal utility as

$$\rho U = (1-\bar{\omega}) \ln\left(\frac{\mu}{(1-\bar{\omega})\mu + \bar{\omega}} \frac{L-N}{La}\right) + \bar{\omega} \ln\left(\frac{1}{(1-\bar{\omega})\mu + \bar{\omega}} \frac{L-N}{La}\right) + \frac{\ln \gamma}{\rho} \lambda N, \tag{10}$$

or

$$\rho U = \ln\left(\frac{\mu^{(1-\bar{\omega})}}{(1-\bar{\omega})\mu + \bar{\omega}}\right) + \ln\left(\frac{L-N}{La}\right) + \frac{\ln \gamma}{\rho} \lambda N.$$

In other words, $\rho U = \Omega - D$ where

$$D \equiv -\ln\left(\frac{\mu^{(1-\bar{\omega})}}{(1-\bar{\omega})\mu+\bar{\omega}}\right) \geq 0 \tag{11}$$

and

$$\Omega \equiv \ln\left(\frac{L-N}{La}\right) + \frac{\ln\gamma}{\rho}\lambda N. \tag{12}$$

D represents the static inefficiency associated with markup μ in the high-technology sector. With Cobb-Douglas utility, the static inefficiency is independent of the level of consumption. Ω represents dynamic social welfare resulting from the allocation of labor between R&D and consumption. Notice that Ω is maximized at N^*. Writing constrained welfare in this way makes clear the trade-off between dynamic and static efficiency. Patents stimulate R&D by allowing successful firms to earn market profits. Stronger patent protection may imply increased R&D, and therefore increased dynamic social welfare when $N < N^*$. However, stronger patent protection also implies larger markups, and therefore increased static inefficiency. In other words, patents can increase growth only at the cost of higher static inefficiencies. Optimal patent policy must weigh this trade-off.

3. Protection Against Future Innovators

In the previous section, we showed that under the endogenous-growth patent policy ψ_o, firms choose suboptimally small innovation size and might hire too little or too much R&D labor. In this section, we ask how different patent policies can correct incentives for R&D. Recall that a specific patent policy is $\psi = (P, K)$, where P is the patentability requirement and K is leading breadth. We now define some notation to denote the equilibrium outcome as a function of patent policy. Let $\hat{\gamma}(\psi)$ be the equilibrium innovation size under policy ψ, $\hat{N}(\psi)$ be the equilibrium level of aggregate R&D labor under policy ψ, $\hat{n}(\psi)$ be the equilibrium level of R&D labor per industry under policy ψ, and $\hat{\phi}(\psi)$ be the equilibrium industry arrival rate under policy ψ. In the previous section, we examined the endogenous-growth patent policy $\psi_o = (P = 1, K = 1)$. Converting notation, $\hat{\gamma}(\psi_o) = \gamma_o$, $\hat{N}(\psi_o) = N_o$, $\hat{n}(\psi_o) = n_o$, and $\hat{\phi}(\psi_o) = \lambda(\gamma_o)n_o$.

Although we found that there can be underinvestment or overinvestment under policy ψ_o, our focus will be the case of underinvestment. We do not examine the case of overinvestment for two reasons. First, there is evidence that suggests there is too little R&D (see in particular Jones and Williams, 1998, 2000), so the overinvestment case is perhaps empirically less relevant. Second, if there is overinvestment under ψ_o, correcting incentives is somewhat trivial. Clearly, ψ_o is not the weakest patent policy. In particular, shorter patent life and/or weaker lagging breadth will decrease the reward to success and therefore reduce the level of R&D. The question arises which technique is better, but this question has been well-studied by Klemperer (1990) and Gilbert and Shapiro (1990) (although in the framework of isolated innovation).

Hence, we now suppose there is underinvestment under policy ψ_o, and ask how protection against future innovators, in the form of a patentability requirement or leading breadth, can stimulate R&D relative to policy ψ_o. Before we discuss the use of patent policy to increase R&D incentives, however, we discuss two caveats to our analysis.

The first caveat revolves around the use of subsidies.[10] In the endogenous-growth literature, the limited attention to how policy can affect the incentive to innovate has focussed on R&D subsidies (and taxes). For example, Grossman and Helpman (1991) and Stokey (1995) correctly point out that an R&D subsidy or tax can induce the first-best level of R&D. If policymakers can use R&D subsidies, however, there is no reason for patents. In fact, if there is a noninnovative sector so markups create output distortions, optimal policy would involve essentially no patent protection and a very large R&D subsidy.[11] In more practical terms, R&D subsidies may be inferior to patents due to asymmetric information between policymakers and firms. Marginal subsidies (or taxes) applied to R&D costs become problematic if firms have flexibility in what they claim to be an ''R&D cost''; patents, in contrast, are granted only if a firm actually achieves an innovation. Lump-sum subsidies for successful firms (i.e., prizes) become problematic if policymakers cannot observe the value of each innovation (see Wright, 1983); in such situations, patents can be useful as a revelation device (see Cornelli and Schankerman, 1999 and Scotchmer, 1999).[12] In sum, there can be problems with R&D subsidies that are less prevalent with patents. Because patents are used extensively and the use of R&D subsidies is limited, we take the perspective that R&D subsidies are not an effective policy instrument.

The second caveat revolves around trade secrets. In general, patent policy matters only to the extent that firms prefer to use patents rather than trade secrecy to protect their intellectual property. In terms of the endogenous-growth literature, if one interprets the usual assumptions not as firms using patents but rather as firms using trade secrets, then the patent policies we suggest below are (most likely) irrelevant. But if we interpret the usual assumptions as firms actually preferring patents to trade secrets—which is certainly implicit in the language used—then the policies discussed below might be useful. A closely related question, however, is whether the policies we discuss induce firms to switch to trade secrets. On one dimension, the answer would appear to be no, because, as will become clear, the policies we discuss increase the value of patents to firms. But a second dimension is, when other firms are using patents, would one firm have an incentive to switch to trade secrecy. While it is hard to answer this question without a formal model of trade secrets, we note that, from a policy-design perspective, we can make the answer no by enforcing leading breadth. In other words, if we enforce leading breadth, then while a firm may be able to use trade secrecy to protect its invention, it would not be able to profit from its invention. In this paper, we focus on the case where either trade secrecy is inherently unattractive or patent enforcement is sufficient to make it unattractive.

10 See Aghion and Howitt (1998) for further discussion of policies that subsidize R&D.

11 Segerstrom (1992) discusses such a policy in the context of uncertain innovation size. With certain innovation size, he shows that if policymakers can also tax the noninnovative sector, they can counteract the markups created by patents. However, an output tax on the noninnovative sector seems to be out of the realm of R&D policy.

12 See also Kremer (1998), who proposes a mechanism that combines R&D subsidies and patents with the goal of letting the marketplace reveal the value of an innovation but then having the government purchase the patent and make the innovation freely available.

3.1. A Patentability Requirement

In a partial-equilibrium framework, O'Donoghue (1998) and Hunt (1999) show that a patentability requirement can stimulate R&D. A patentability requirement induces firms to pursue larger innovations which take longer to achieve. Hence, a successful firm earns market profits for a longer period of time. This increases the reward to success and therefore stimulates R&D. We now demonstrate that this result can hold in the endogenous-growth model.

When there is a patentability requirement $P \geq \gamma_o$ and no leading breadth, the outcome of the model is exactly analogous to the outcome under policy ψ_o. The only difference is that firms will choose innovation size P because P represents a lower bound on the innovation size that firms will pursue (assuming firms want to obtain a patent). The following lemma is analogous to Lemma 1:

Lemma 2 For any policy $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = 1$:

i. The market leader in industry $\omega \in [0, \bar{\omega}]$ does not conduct R&D.

ii. Firms choose innovation size $\hat{\gamma}(\psi) = P$.

iii. The markup in industry $\omega \in [0, \bar{\omega}]$ is $\mu = \hat{\gamma}(\psi)$.

iv. The equilibrium level of R&D is

$$\hat{N}(\psi) = \bar{\omega}\hat{n}(\psi) = \begin{cases} \bar{\omega}\left[\frac{P-1}{P}L - \frac{(1-\bar{\omega})P+\bar{\omega}}{P}\frac{\rho}{\lambda(P)}\right] & \text{if } L \geq \frac{(1-\bar{\omega})P+\bar{\omega}}{P-1}\frac{\rho}{\lambda(P)} \\ 0 & \text{if } L \leq \frac{(1-\bar{\omega})P+\bar{\omega}}{P-1}\frac{\rho}{\lambda(P)}. \end{cases}$$

Proposition 1 establishes that imposing a patentability requirement can stimulate R&D in the endogenous-growth framework (all proofs are collected in Appendix B).

Proposition 1 *(A patentability requirement can stimulate R&D.) Consider a policy $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = 1$:*

i. Suppose $N_o > 0$. Then there exists $P' > P'' > \gamma_o$ such that $\hat{N}(\psi)$ is increasing in P for all $P \in [\gamma_o, P')$ and $\hat{\phi}(\psi)$ is increasing in P for all $P \in [\gamma_o, P'')$.

ii. Suppose $N_o = 0$. Then either (a) $\hat{N}(\psi) = \hat{\phi}(\psi) = 0$ for all $P \geq \gamma_o$; or (b) there exists $P' > P'' > \underline{P} \geq \gamma_o$ such that $\hat{N}(\psi) = \hat{\phi}(\psi) = 0$ for all $P \in [\gamma_o, \underline{P}]$, $\hat{N}(\psi)$ is increasing in P for $P \in [\underline{P}, P')$, and $\hat{\phi}(\psi)$ is increasing in P for $P \in [\underline{P}, P'')$.

Part (i) of Proposition 1 states that as long as there is investment under policy ψ_o, imposing a patentability requirement can induce firms to hire more R&D labor, and that doing so can lead to an increased rate of innovation. Part (ii) then states that if there is no investment under policy ψ_o, imposing a patentability requirement may stimulate investment; however, it is possible that investment might not occur for any patentability requirement.

Proposition 1 by itself does not establish whether a patentability requirement is welfare enhancing. Imposing a patentability requirement $P \geq \gamma_o$ has three effects on social welfare. First, a patentability requirement induces firms to pursue larger innovations. Second, a patentability requirement induces firms to hire more R&D labor (at least in the relevant range). Since under ψ_o firms pursue suboptimally small innovations, in the case where there is underinvestment under ψ_o these two effects will both lead to increased dynamic social welfare. However, the third effect is that a patentability requirement creates increased industry markups, and therefore decreased static efficiency. The optimal patentability requirement must weigh the trade-off between dynamic and static efficiency.

3.2. Leading Breadth

An alternative way to stimulate R&D is proposed by O'Donoghue et al. (1998). Again in a partial-equilibrium framework, they show that leading breadth can stimulate R&D by allowing firms to consolidate market power through licensing agreements. To illustrate this intuition, consider the policy $\psi = (P = \gamma_o, K = (\gamma_o)^2)$, and assume for the time-being that all firms choose innovation size γ_o. Suppose the most-recent innovator has a patent on quality q_o. Quality q_o infringes the patent of the second-most-recent innovator (who has a patent on quality q_o/γ_o), but does not infringe the patent of the third-most-recent innovator (who has a patent on quality $q_o/(\gamma_o)^2$). Hence, in order for the maximum feasible quality q_o to be produced, the most-recent innovator and the second-most-recent innovator must enter a licensing agreement (because neither is legally allowed to produce quality q_o without violating the other's patent). When they do so, their nearest competitor will be the third-most-recent innovator, and therefore the two firms in the licensing contract can charge markup $(\gamma_o)^2$, compared to markup γ_o under policy ψ_o. In other words, leading breadth facilitates collusion between the most-recent and second-most-recent innovators by forcing them to negotiate over permission to produce. Of course, the exact reward to success will depend on how profit is shared in these licensing agreements.

Before proceeding to our formal analysis of such policies, we discuss their real-world applicability. A first question is whether it is reasonable to expect the patent authorities to allow firms to patent an invention when that invention infringes on some prior patent. In fact, such ''blocking'' patents are very much a part of existing patent law. Blocking patents arise when an invention represents an improvement over some prior technology that is judged to satisfy the requirements to receive a patent but not to be extensive enough to be free from the prior patent. A firm that receives a blocking patent cannot use its new technology without permission from the prior infringed patentholder; however, the value of the blocking patent comes from the fact that the prior patentholder also cannot use the new technology without permission from the holder of the blocking patent.[13]

A second question is how do real-world firms react when the frontier technology

13 For more discussion of blocking patents, see Merges and Nelson (1990, 860–862) and Gilbert (2002, 6–7). The existence of blocking patents illustrates how patents are not a right to use a technology, but rather are a right to exclude others from using a technology (see footnote 96 in Merges and Nelson).

infringes on multiple patents. Not surprisingly, the answer differs across industries and across time. But one particular response is becoming more common: ''patent pools''.[14] A patent pool is an agreement among multiple patentholders to pool their patents, often by creating a central administrative entity. The pool specifies licensing rates for access to the entire pool of patents, and it also specifies a division of proceeds among the pool's members. As an example, in 1995 nine patentholders agreed to pool the 27 patents that were required to use the MPEG-2 video compression technology. The pool was set up with a mechanism for evaluating whether new technologies should be added to the pool and whether older technologies should be removed from the pool, and with a mechanism for recalibrating the division of proceeds when such changes are made.

A third question is what has been the antitrust reaction to patent pools. Historically, patent pools have existed since the mid-19th century (Merges (1999) notes that the first patent pool appeared in 1856 in the sewing machine industry). But patent pools became relatively scarce from the 1940s until the 1990s, in large part due to the negative antitrust reaction to any industry-wide agreements. A negative antitrust response is justified to some degree, because patent pools are a natural way to mask cartel agreements—in the early part of this century, many patent pools were clearly such. In the 1990s, however, the antitrust authorities recognized that legitimate patent pools may promote social welfare, both by facilitating the use of frontier technologies, and by increasing incentives to produce new technologies. As a result, the use of patent pools has grown in recent years.

With this real-world background, we now analyze a stylized model of the effects of leading breadth—which determines the extent of blocking relationships—when the antitrust authorities permit licensing agreements to incorporate only the owners of essential patents (such licensing agreements might take the form of a patent pool). The key question concerning leading breadth is therefore the number of patents that the maximum feasible quality infringes. To simplify our analysis, we focus on policies $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = P^\alpha$ for some $\alpha \in \{1, 2, \ldots\}$. As we show below, under such policies, firms target innovation size P, and therefore the maximum feasible quality infringes α patents (its own patent and $\alpha - 1$ previous patents).[15]

The reward to success will depend on how profit is shared in licensing agreements. O'Donoghue et al. (1998) avoid dealing with the division of profit by assuming firms can bargain over R&D costs. They focus on a model without patent races, however, and with patent races bargaining over R&D costs seems inappropriate. Indeed, the basic premise of the patent race model is that there is nothing to bargain over until some firm has a success,

---

14 This discussion draws heavily from Merges (1999) and Gilbert (2002).

15 Our conclusion that firms target innovation size P is driven by our restriction to policies with $K = P^\alpha$. For more general policies, firms need not target innovation size P—for example, if $P = \gamma'$ and $K = \gamma' + \varepsilon$ for some small $\varepsilon > 0$, firms might target innovation size K so as to avoid the leading breadth. Even so, from the perspective of optimal policy design, our restriction to policies with $K = P^\alpha$ is not restrictive. In our model, P and K matter only insofar as they influence (1) the innovation size γ that firms pursue and (2) the number of patents α that the maximum feasible quality infringes. For any γ and α, the most direct way to implement that outcome is to set $P = \gamma$ and $K = P^\alpha$. If, for instance, under policy $P = \gamma'$ and $K = \gamma' + \varepsilon$ firms target innovation size K and therefore $\alpha = 1$, the same outcome could be implemented by setting $P = \gamma' + \varepsilon$ and $K = P$.

but at that point the R&D costs have already been sunk. Even so, we do not wish to explore here the intricacies of licensing. Hence, we use a vastly simplified model of licensing to make the analysis tractable and to provide some intuition as to how leading breadth affects the incentive to innovate.

We simplify our analysis by assuming that firms who participate in the licensing contract do not conduct R&D.[16] Hence, under policy $\psi = (P, K)$ with $K = P^\alpha$, exactly α patentholders will participate in the licensing arrangement. These firms will split market profits according to some bargaining solution. We suppose there is a set of stationary bargaining solutions $(\mathbf{s}^1, \mathbf{s}^2, \ldots)$ where for each $\alpha \in \{1, 2, \ldots\}$, $\mathbf{s}^\alpha \equiv (s_1^\alpha, s_2^\alpha, \ldots, s_\alpha^\alpha) \in [0, 1]^\alpha$ and $\Sigma_{i=1}^{\alpha} s_i^\alpha = 1$. If α patentholders sign the licensing contract, the most-recent innovator gets share s_1^α, the second-most-recent innovator gets share s_2^α, and so on. In addition, we assume for simplicity that the bargaining solution is the same in all industries.

Although we have not yet defined an equilibrium when there is leading breadth, we can describe some features that must hold in any outcome with this simplified model of licensing.

Lemma 3 Suppose the patent policy is $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = P^\alpha$. Conditional on innovation rate ϕ, any equilibrium will be characterized by the following features:

i. Firms in industry $\omega \in [0, \bar{\omega}]$ target innovation size $\hat{\gamma}(\psi) = P$,

ii. The markup in industry $\omega \in [0, \bar{\omega}]$ is $\mu = P^\alpha$, and

iii. The reward to success in industry $\omega \in [0, \bar{\omega}]$ is $V = \pi \cdot B(\phi, \alpha)$, where π is given by equation 8 and $B(\phi, \alpha) \equiv \Sigma_{i=1}^{\alpha} s_i^\alpha \phi^{i-1}/(\rho + \phi)^i$.

When there is no leading breadth (i.e., $\alpha = 1$), any patentability requirement $P \geq \gamma_o$ will be binding. Part (i) of Lemma 3 establishes that even for the case where $\alpha > 1$, any patentability requirement $P \geq \gamma_o$ will be binding. Choosing innovation size $\gamma > P$ to influence marginal profits is unattractive for much the same reason as it is when $\alpha = 1$. A second reason to choose $\gamma > P$ here is to avoid the leading breadth of some prior patent, but our restriction to $K = P^\alpha$ makes this prohibitively costly. Part (ii) of Lemma 3 simply reflects the fact that increased leading breadth creates increased market profits (as shown by O'Donoghue et al. 1998). Specifically, the markup μ is increasing in the amount of leading breadth (as parametrized by α), and market profits are increasing in the markup μ. Part (iii) of Lemma 3 illustrates that in addition to market profits, the reward to success depends on the bargaining solution. We have defined $B(\phi, \alpha) = \Sigma_{i=1}^{\alpha} s_i^\alpha \phi^{i-1}/(\rho + \phi)^i$ to be the discounted share of market profits that each innovator receives. We often refer to $B(\phi, \alpha)$ as the bargaining discount factor.

The magnitude of $B(\phi, \alpha)$ depends on the bargaining solution $\mathbf{s}^\alpha$, and in particular on

---

16 Because there are constant returns to R&D spending, no inefficiencies are created when the market leader and the other firms in the licensing contract do not conduct R&D.

whether payoffs tend to be frontloaded vs. backloaded. For any α, $B(\phi, \alpha)$ is maximized under bargaining solution $\mathbf{s}^\alpha = (1, 0, \ldots, 0)$ when $B(\phi, \alpha) = 1/(\rho + \phi)$, and $B(\phi, \alpha)$ is minimized under bargaining solution $\mathbf{s}^\alpha = (0, 0, \ldots, 1)$ when $B(\phi, \alpha) = \phi^{\alpha-1}/(\rho + \phi)^\alpha$. Because an innovator's bargaining position is weakest early in the life of her patent, when her product infringes previous patents, and strongest late in the life of her patent, when previous patents have effectively expired and subsequent innovations now infringe her patent, it seems likely that payoffs would be to some extent backloaded. Of course, the backloading of payoffs matters only because firms discount the future (i.e., $\rho > 0$)—for any $\mathbf{s}^\alpha$, $\lim_{\rho\to 0} B(\phi, \alpha) = 1/\phi$.

In addition to the conditions in Lemma 3, an equilibrium with leading breadth must satisfy the no-profit condition $\lambda V = w$, which we can rewrite as $\lambda \pi B(\phi, \alpha) = w$. When there is leading breadth and $\alpha \geq 2$, however, there can be multiple industry arrival rates ϕ that satisfy this no-profit condition. See Appendix A for a more complete discussion. We define the equilibrium to be the largest industry arrival rate that satisfies the no-profit condition.[17] In other words,

$$\hat{\phi}(\psi) \equiv \begin{cases} \max\{\phi \mid \lambda V = w\} & \text{if the set } \{\phi \mid \lambda V = w\} \text{ is non-empty} \\ 0 & \text{if the set } \{\phi \mid \lambda V = w\} \text{ is empty.} \end{cases}$$

Now consider how introducing leading breadth influences the behavior of R&D firms. Increased leading breadth has two effects on the reward to success $V = \pi\, B(\phi, \alpha)$. First, increased leading breadth leads to a larger markup μ and therefore increased market profits π. This effect will tend to stimulate R&D spending. Increased leading breadth also changes the bargaining discount factor $B(\phi, \alpha)$. If stronger leading breadth increases $B(\phi, \alpha)$, then this second effect will also stimulate R&D spending. But it is perhaps more likely that stronger leading breadth reduces $B(\phi, \alpha)$—indeed, $B(\phi, 1) \geq B(\phi, \alpha)$ for all ϕ and $\alpha > 1$. Intuitively, stronger leading breadth can make payoffs more backloaded, because any given innovation infringes more previous patents and therefore is in an even weaker bargaining position early in the life of its patent. Proposition 2 establishes that leading breadth can stimulate R&D as long as increased leading breadth does not cause the bargaining solution to become excessively backloaded.

Proposition 2 *(Leading breadth can stimulate R&D.) Consider two policies $\psi = (P, K)$ and $\psi' = (P, K')$ with $P \geq \gamma_o$ and $K' > K$, and suppose $\hat{N}(\psi) > 0$.*

i. Suppose $K = P$ and $K' = P^2$. If the bargaining solution $\mathbf{s}^2 \equiv (s, 1-s)$, then there exists $\bar{s} \leq P/(P+1)$ such that $\hat{N}(\psi') \geq \hat{N}(\psi)$ as long as $s \geq \bar{s}$.

ii. Suppose $K = P^\alpha$ and $K' = P^{\alpha+1}$. For any bargaining solutions $\mathbf{s}^\alpha$ and $\mathbf{s}^{\alpha+1}$ there exists $\overline{\rho} > 0$ such that $\hat{N}(\psi') \geq \hat{N}(\psi)$ as long as $\rho < \overline{\rho}$.

17 We show in Appendix A that this definition is well-defined.

Part (i) of Proposition 2 establishes that, relative to policy ψ_o, increasing leading breadth so that $\alpha = 2$ will stimulate R&D as long as the first-period share of market profits is big enough. Part (ii) of Proposition 2 establishes that increased leading breadth can always stimulate R&D investment as long as people are patient enough. Proposition 2 illustrates that one of two things must be true for the profit effect of increased leading breadth to dominate the backloading effect. First, it could be that increased leading breadth does not significantly increase backloading (as illustrated by part (i)). Second, it could be that people are patient enough so that backloading is not too costly (as illustrated by part (ii)).[18]

It is worth reiterating that the mechanism behind leading breadth stimulating R&D is that leading breadth facilitates collusion. If leading breadth did not facilitate collusion—for example, if the antitrust authorities restricted the members of the licensing agreement to charge a markup no larger than the most-recent innovation—then leading breadth would retard R&D. Intuitively, in this case, total (undiscounted) market profits are unchanged, but firms receive these profits with more delay. Hence, the effectiveness of leading breadth depends on some coordination between the patent authorities and the antitrust authorities (and as alluded to earlier, there has been more and more coordination since the early 1990s).[19]

Just as Proposition 1 does not establish whether a patentability requirement is welfare enhancing, Proposition 2 does not establish whether leading breadth is welfare enhancing. Imposing leading breadth has two effects on social welfare. First, leading breadth can induce firms to hire more R&D labor. In the case where there is underinvestment under ψ_o this effect will lead to increased dynamic social welfare. Second, leading breadth allows firms to consolidate market power and thereby creates increased industry markups. This effect will lead to decreased static efficiency. Like the optimal patentability requirement, optimal leading breadth must weigh the trade-off between dynamic and static efficiency.

How does a patentability requirement compare to leading breadth? Since leading breadth allows firms to consolidate market power and create large market profits, it can be much more effective than a patentability requirement at stimulating R&D. However, these large market profits may make leading breadth significantly worse than a patentability requirement in terms of static efficiency, and in addition the effectiveness of leading breadth can be undermined by licensing inefficiencies. A careful comparison of the two types of policy would require much further specification of the model, and is beyond the scope of this paper. Furthermore, the appropriate question is not whether a patentability requirement or leading breadth is better, but rather is how to tailor the two instruments together to improve incentives for R&D. Indeed, Proposition 2 is written in a way that illustrates how leading breadth can be effective in addition to a patentability requirement.[20]

---

18 It is straightforward to show that there also exists $\bar{\rho}' > 0$ such that there is underinvestment under policy ψ_o for all $\rho < \bar{\rho}'$. Hence, for ρ small there is underinvestment under ψ_o and leading breadth can stimulate R&D.

19 A more subtle question is what the effect of leading breadth would be if there were already collusion in the absence of leading breadth. The answer would depend on the form of collusion in the absence of leading breadth—what is the markup and what are the relative bargaining positions—and how the presence of leading breadth influences that collusion. An analysis of this issue is beyond the scope of this paper; however, it is clearly an important consideration for any real-world application of leading breadth.

20 For more discussion of these issues, see O'Donoghue (1998).

4. The Role of Scale Effects

Our analysis above demonstrates how protection against future innovators, in the form of a patentability requirement or leading breadth, might be a useful method to encourage industrial R&D and economic growth. To lay bare the basic intuitions, however, we have embedded our analysis within the simple quality-ladder model of Grossman and Helpman (1991). This model and other first-generation R&D-based growth models (e.g., Romer, 1990; Segerstrom et al. 1990; Aghion and Howitt, 1992) have been criticized on the basis that they exhibit ''scale effects'', wherein the economy-wide growth rate of the knowledge stock is proportional to resources employed in the R&D sector. Jones (1995) rejects these models on empirical grounds because such scale effects are counterfactual: In almost all industrial countries, the post-war period was characterized by a rapid increase in research employment without a corresponding increase in the growth rate of total factor productivity.

In this section, we address whether our main conclusions would survive in the more elaborate second- and third-generation growth models that do not exhibit scale effects. We divide our discussion into two dimensions. First, under various alternative models that remove the scale effect, are our policies still effective at influencing the allocation of labor to R&D? Second, when our policies are effective at influencing the allocation of labor to R&D, what are the implications for economic growth?

We begin with the latter question. Following Jones (1999), growth models that do not exhibit the scale effect can be divided into second-generation and third-generation models. Second-generation models remove the scale effect by assuming that innovations become more difficult to achieve as the technological frontier expands. In Jones' notation, the technological frontier (the stock of knowledge) A develops according to $\dot{A}/A = \delta L_A A^{\phi - 1}$, where $\phi < 1$ and δ are constants, and L_A is total labor allocated to R&D. If ϕ is sufficiently small, $\dot{A}/A$ may stagnate (or decrease) even as the amount of resources in R&D L_A is growing. In such environments, long-run (steady-state) growth rates are independent of the level of R&D resources, and therefore even if our policies are effective at stimulating R&D, they cannot affect long-run growth rates. However, they would still increase rate of convergence, which might be welfare-enhancing.

Third-generation models remove the scale effect by endogenizing the number of product varieties—so there is horizontal R&D that invents new varieties and vertical R&D that improves existing varieties—and by assuming that the growth rate depends on research effort per variety. Because larger populations are associated with more varieties, there may be more workers doing research without an increase in economic growth as these R&D workers are spread across more varieties. Such models restore the basic conclusion from first-generation models that the allocation of labor influences steady-state growth rates. Hence, in such models, if our policies are effective at stimulating R&D spending per variety, they may increase steady-state growth rates.

With this background in mind, we now discuss whether our policies can be effective at influencing the allocation of labor to R&D in various second- and third-generation growth models. The key question here is whether the structure of the particular model permits us to introduce our policies in a meaningful way. In particular, the logic behind our results is that a patentability requirement and leading breadth can be useful when R&D takes place within product lines and when there is leapfrogging (market leaders are repeatedly replaced by innovators), because these policies are designed to protect against the threat of being displaced as the market leader.

Among second-generation growth models, the seminal paper by Jones (1995) assumes horizontal product innovations, and existing technologies are never rendered obsolete. Clearly, our policies would not be useful in such an environment. However, many second-generation models, including Segerstrom (1998), Kortum (1997), and Li (2000, 2002), assume that R&D takes place within product lines and that there is leapfrogging. In these models, the basic logic of our results should hold, and so a patentability requirement and leading breadth might be useful to increase R&D. Of course, as discussed above, our policies could have only temporary effects on economic growth (on the speed of convergence), and would not affect long-run (steady-state) economic growth.

Most third-generation models build environments in which our policies would not be effective. For instance, Young (1998) assumes that time is discrete and that innovators survive for exactly one period. For obvious reasons, our policies cannot be useful in such a context. Similarly, Smulders and van der Klundert (1995), Peretto (1998a,b), Dinopoulos and Thompson (1998), and Peretto and Smulders (1998) assume that entrants invent new varieties and that, once successful, the firm conducts R&D to lower its production costs and/or improve the quality of its products. Importantly for our purposes, however, these models all assume that the inventor of a specific variety will remain the market leader in that variety forever after. Hence, there is no leapfrogging—no threat from future innovators—and so our policies would not be effective.

Howitt (1999), in contrast, builds a third-generation model with leapfrogging within product lines.[21] In this environment, our policies can be useful for increasing the amount of labor allocated to vertical R&D and thereby influence long-run economic growth. Indeed, in Appendix C we present a modified version of Howitt's model that does not exhibit the scale effect in which our main conclusions hold. In sum, our policies are not useful in most third-generation models, but only because these models have eliminated certain features of market competition—most notably from our perspective, the feature of leapfrogging. Since we suspect the reason for doing so was tractability, and not a belief that leapfrogging is unrealistic, and since our policy conclusions survive in a third-generation model with leapfrogging, we believe our results are quite relevant for the endogenous-growth literature.

5. Partial Equilibrium vs. General Equilibrium

Our goal in Section 3 was to describe what the endogenous-growth literature can learn from the patent-design literature. Our goal in this section is to describe what the patent-design literature can learn from the endogenous-growth literature. By confining itself to partial-equilibrium analyses, the patent-design literature examines how patents affect a single

21 See also Aghion and Howitt (1998, Ch. 12), who suggest a modification of their basic model that permits the number of distinct products to be endogenous.

industry assuming nothing changes elsewhere in the economy. However, patent policy is an economy-wide phenomenon. Patents are used in many industries, and therefore any policy change will affect many industries. The general-equilibrium approach of the endogenous-growth literature therefore permits a more complete analysis of patent design.[22]

An important general-equilibrium effect that is not revealed in partial-equilibrium analyses concerns the impact of patent policy on the real wage. Partial-equilibrium analyses usually take the cost of R&D to be fixed and exogenous, and in particular to be independent of patent policy. But in a general-equilibrium framework in which patent policy applies to many industries, patent policy influences the real wage, and therefore affects the real cost of R&D. Specifically, patent policy operates by increasing profits in high-technology industries, which implies a redistribution of aggregate income from wages to profits, and therefore a decline in the real wage. Moreover, the effectiveness of patent policy depends on by how much high-technology profits increase relative to the wage, and so this general-equilibrium effect reinforces the effects of patents. Because a larger high-technology sector increases the impact of patent policy on the real wage, the effectiveness of patent policy is increasing in the size of the high-technology sector. Proposition 3 captures this intuition by showing that for any specific patent policy, the bigger is the high-technology sector the larger is the equilibrium industry arrival rate induced by that patent policy. Proposition 3 implies that the bigger the high-technology sector, the weaker is the protection needed to implement any target industry innovation rate. (A partial-equilibrium analysis is in a sense equivalent to assuming $\bar{\varpi}$ close to 0.)

Proposition 3 *(Patent policy is more effective as more industries innovate). For any patent policy* $\psi = (P, K)$ *with* $P \geq \gamma_o$ *and* $K = P^{\alpha}$, $\hat{\phi}(\psi)$ *is increasing in* $\bar{\varpi}$.

The economy-wide effects of patents also has implications for static inefficiency. The core theme of the patent-design literature is the trade-off between incentives for R&D and monopoly distortions: Patents create markups that lead to increased incentives to innovate, but these markups cause static inefficiency. However, there is often no formal model of static inefficiency—the usual measure is the deadweight loss associated with the assumed demand curve, but this measure is often invalid in a general-equilibrium framework. The general-equilibrium framework allows us to formally model the static inefficiency. As is the explicit or implicit assumption in much of the patent-design literature, static inefficiency arises in the form of output distortions.[23] The markups created by patents distort consumption away from high-technology industries and towards noninnovative industries. In our model, the static inefficiency is given by equation (11).

The perhaps surprising result is that the social cost of output distortions may be smaller

22 In fact, our discussion of what the general-equilibrium framework contributes to the patent-design literature applies equally well to cumulative innovation and isolated innovation. Our discussion revolves around the cumulative-innovation interpretation only because we focus on a model of endogenous growth.

23 Some patent-design papers concentrate on other static inefficiencies such as delayed diffusion or high-cost firms persisting in the market.

than partial-equilibrium analyses assume. In particular, if $\bar{\omega}$ is close to 0, which means few industries conduct R&D, or if $\bar{\omega}$ is close to 1, which means most industries conduct R&D, then output distortions are not very costly. For $\bar{\omega}$ close to 0, patents create markups in only a few industries, so relative prices are efficient except for the few innovative industries. In this case, output distortions involve too little consumption in the few innovative industries, but very little change in the noninnovative industries. Since the innovative industries produce a small portion of the consumption bundle, these output distortions are not too costly. For $\bar{\omega}$ close to 1, patents create an identical markup in almost all industries, so relative prices are efficient except for the few noninnovative industries. Again, the resultant output distortions are not too costly. Proposition 4 summarizes this discussion.

Proposition 4 *(Output distortions are small for few or many innovative industries). For any policy $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = P^{\alpha}$, there exists $\omega' \in (0, 1)$ such that the static inefficiency D is increasing in $\bar{\omega}$ for all $\bar{\omega} < \omega'$ and decreasing in $\bar{\omega}$ for all $\bar{\omega} > \omega'$. Moreover, for $\bar{\omega} = 0$ and $\bar{\omega} = 1$ there is no static inefficiency.*

We conclude our discussion of partial-equilibrium vs. general-equilibrium with two general-equilibrium effects that are absent from our model. The first revolves around the effects of patent policy on labor supply. Our model assumes that the aggregate supply of labor is exogenously given. If instead we were to assume that labor supply is increasing in the real wage (i.e., the substitution effect is stronger than the income effect), there would be a mitigating force that reduces the effectiveness of patents. Specifically, because stronger patents tend to reduce the real wage, stronger patents would lead to a reduction in the supply of labor. This reduction in the supply of labor would mitigate the effects of stronger patent policy. Elastic labor supply would also imply that, in contrast to Proposition 4, patents might create static inefficiency even when most industries are innovative, in the form of reduced labor supply.

The second revolves around the effects of patent policy on the allocation of demand across industries. Our model assumes that consumers have homothetic preferences, and so changes in the distribution of income do not have feedback effects on the allocation of demand. If, however, income elasticities were not equal to unity for all products, stronger patents might lead to a reallocation of demand, which in turn might have feedback effects on the incentive to conduct R&D.[24]

6. R&D Distortions

An important issue largely ignored by both the endogenous-growth literature and the patent-design literature is the allocation of R&D across industries. The endogenous-

24 Of course, models of innovation and growth that feature a balanced-growth path are not easily conducive to the analysis of non-homothetic preferences. For models where growth is driven by vertical R&D and where income distribution feeds back to the demand of innovators, see Li (2002), Glass (2001), and Zweimüller and Brunner (1998).

growth assumption of identical R&D capabilities in all innovative industries implies that the allocation of R&D labor is irrelevant. As a result, the only policy concern in the previous two sections was the allocation of labor between R&D and production. This assumption, however, is highly questionable. Empirical evidence suggests there are significant cross-industry differences in R&D productivity and in R&D behavior (see the survey by Cohen and Levin, 1989). Klenow (1996) builds an endogenous-growth model in which he allows industries to vary along several dimensions in order to explore the source of these cross-industry differences. Comparing the implications of his model to the empirical evidence, he concludes that industry differences in market size and technological opportunities best explain industry differences in R&D behavior.

In this section, we posit that there are asymmetric R&D capabilities across industries (i.e., asymmetric technological opportunities), and ask what are the implications for patent policy. For simplicity, we shall assume that all industries are innovative (i.e., $\bar{\omega} = 1$), and that a uniform patent policy applies for all industries.

We suppose asymmetric R&D capabilities arise from sector-specific capital. (It will become clear that our basic results would hold for other sources of asymmetric R&D capabilities.) Suppose capital is used only in R&D. If an R&D firm employs labor n and capital h, it will have arrival rate of innovations $\phi \equiv \lambda(\gamma)h^{1-\beta}n^{\beta}$, $\beta \in (0, 1)$. The function λ is exactly as in the basic model, and in the discussion that follows we often suppress its argument. There are two types of capital, sector-1 capital and sector-2 capital, where all industries with $\omega \in [0, \frac{1}{2}]$ can use only sector-1 capital and industries with $\omega \in (\frac{1}{2}, 1]$ can use only sector-2 capital. All results easily generalize to the case where the two sectors are of unequal sizes, but the notation would be more complicated. The total supply of sector-1 capital is H_1, the total supply of sector-2 capital is H_2, and all consumers own equal shares of each type of capital (which is part of their wealth $A(t)$). The remainder of the model is unchanged.

With two sectors conducting R&D, we need some new notation. Specifically, we let N_1 denote total labor employed in sector-1 R&D, N_2 denote total labor employed in sector-2 R&D, and N denote the aggregate level of labor employed in R&D, so $N \equiv N_1 + N_2$.

6.1. The Social Optimum

A social planner now must allocate labor between production, sector-1 R&D, and sector-2 R&D—that is, a social planner must choose x_ω for each $\omega \in [0, 1]$, N_1, and N_2. Since there are constant returns to scale in R&D, for any N_1 a social planner will allocate sector-1 R&D resources in the ratio H_1/N_1 (i.e., all sector-1 firms that conduct R&D will receive capital h and labor n such that $h/n = H_1/N_1$). As a result, the total arrival rate of innovations in sector 1 is $\lambda H_1^{1-\beta} N_1^{\beta}$. Analogously, for any N_2 a social planner will allocate sector-2 R&D resources in the ratio H_2/N_2, and the total arrival rate of innovations in sector 2 is $\lambda H_2^{1-\beta} N_2^{\beta}$. Using a logic similar to that in Section 2, we can derive the following expression for intertemporal utility:

$$\rho U = \ln\left(\frac{L - (N_1 + N_2)}{La}\right) + \frac{\ln \gamma}{\rho} \lambda(\gamma) \left[H_1^{1-\beta} N_1^{\beta} + H_2^{1-\beta} N_2^{\beta}\right]. \tag{13}$$

A social planner will choose the endogenous variables N_1, N_2, and γ to maximize equation (13). The socially optimal innovation size is the same γ^* as in the basic model. The social planner has two concerns when allocating labor: (1) how much labor to allocate to R&D (i.e., the choice of N), and (2) how to divide R&D labor between sector-1 industries and sector-2 industries (i.e., the choice of N_1/N_2). Our focus in this section is the latter decision. It is straightforward to show that the optimal distribution of R&D labor is $[N_1/N_2]^* = H_1/H_2$. Moreover, conditional on any N and γ, intertemporal utility is maximized when $N_1/N_2 = [N_1/N_2]^*$.

6.2. The Private Equilibrium

The only feature of the economy different from the basic model is the R&D production function. The demand functions for consumption goods, optimal firm behavior in output markets, and the resource constraint (for labor) are all unchanged. We again incorporate these features into equation (8), and since $\bar{\omega} = 1$ we have

$$\bar{\pi} = (L - N)(\mu - 1), \tag{14}$$

where μ is the industry markup determined by patent policy. Since there is a uniform patent policy economy-wide, the markup μ will be the same in all industries, and therefore market profits will be the same in all industries.

The introduction of sector-specific capital creates important changes in a firms' R&D decisions. As in the basic model, there is perfect competition in R&D and the R&D production function has constant returns to scale. As a result, the individual research venture is of indeterminate size. However, since all industries within a given sector are symmetric, the amount of capital and labor employed in R&D at the industry level will be the same across all industries within a given sector. In other words, there will exist h_1, n_1, h_2, and n_2 such that every industry in sector 1 will employ capital h_1 and labor n_1 and every industry in sector 2 will employ capital h_2 and labor n_2. Hence, every industry in sector 1 will have industry arrival rate $\phi_1 = \lambda h_1^{1-\beta} n_1^{\beta}$, and every industry in sector 2 will have industry arrival rate $\phi_2 = \lambda h_2^{1-\beta} n_2^{\beta}$. Also, because industries with $\omega \in [0, \frac{1}{2}]$ are sector-1 industries and industries with $\omega \in (\frac{1}{2}, 0]$ are sector-2 industries, $N_1 = \frac{1}{2} n_1$, $N_2 = \frac{1}{2} n_2$, $H_1 = \frac{1}{2} h_1$, and $H_2 = \frac{1}{2} h_2$.

The introduction of sector-specific capital changes the reward to success only in the sense that sector-1 industries and sector-2 industries have different arrival rates and therefore different rewards to success. Specifically, the reward to success for each industry in sector j is

$$V_j = \bar{\pi} \cdot B(\phi_j, \alpha) = \bar{\pi} \sum_{i=1}^{\alpha} s_i^{\alpha} \frac{\phi_j^{i-1}}{(\rho + \phi_j)^i}. \tag{15}$$

Now consider an individual firm's decision. Let w_j be the rental rate for sector-j capital. If a firm in sector j employs labor n and capital h, the net payoff to R&D is

$$-n\,w - h\,w_j + \lambda\, h^{1-\beta} n^{\beta} V_j.$$

The choice of innovation size is identical to that in the basic model—in particular, under policy ψ_o firms choose innovation size γ_o, and under any policy with patentability requirement $P \geq \gamma_o$ firms choose innovation size *P*. In terms of capital and labor, profit maximization implies that all firms in sector *j* employ capital *h* and labor *n* in the same ratio, which must be $h/n = h_j/n_j$. Perfect competition in capital markets implies that the rental rate for sector-*j* capital is equal to its marginal product, or $w_j = (1-\beta)\lambda h_j^{-\beta} n_j^{\beta} V_j$. Using $h/n = h_j/n_j$ and $w_j = (1-\beta)\lambda h_j^{-\beta} n_j^{\beta} V_j$, we can write the no-profit conditions for the two sectors as

$$\beta\lambda\left(\frac{h_1}{n_1}\right)^{1-\beta} V_1 = w \tag{16}$$

$$\beta\lambda\left(\frac{h_2}{n_2}\right)^{1-\beta} V_2 = w. \tag{17}$$

Equations (14)–(17) determine the equilibrium allocation of labor. However, as in the basic model, whereas for any patent policy ψ with $\alpha = 1$ there is a unique outcome that satisfies these no-profit conditions, for any patent policy ψ with $\alpha \geq 2$ there can be multiple outcomes that satisfy these no-profit conditions. As before, we define the equilibrium allocation to be that with the largest arrival rates that satisfy the no-profit condition. See Appendix A for a more complete discussion.

6.3. *The Role of Patents*

We focus on the following two questions: What R&D distortions are associated with the use of patents? How do R&D distortions depend on the strength of patent protection?

Given R&D production function $\phi = \lambda h^{1-\beta} n^{\beta}$, R&D labor is more productive in the sector with more available sector-specific capital per industry. For the remainder of this section, we assume without loss of generality that $h_1 > h_2$, so R&D labor is more productive in sector-1 industries. Appendix A establishes that for any patent policy ψ, $h_1 > h_2$ implies $\hat{\phi}_1(\psi) > \hat{\phi}_2(\psi)$—the sector in which R&D labor is more productive will always have a larger industry arrival rate.

Consider first policy $\psi_o = (P = 1, K = 1)$. It is straightforward to show that under policy ψ_o firms pursue suboptimally small innovations. It is also straightforward to show that under ψ_o there can be too little or too much aggregate labor allocated to R&D. As emphasized above, however, our focus in this section is the allocation of R&D labor between sectors 1 and 2. The following proposition establishes that under policy ψ_o the private equilibrium allocates too little R&D labor to the more productive sector.

Proposition 5 *(Patents distort R&D labor away from productive industries.) Suppose* $h_1 > h_2$*, so R&D labor is more productive in sector-1 industries. Under policy* ψ_o*,* $\hat{N}_1(\psi_o)/\hat{N}_2(\psi_o) < [N_1/N_2]^*$.

Proposition 5 is driven by differential rates of creative destruction in the two sectors. Since R&D labor is more productive in sector-1 industries, these industries will have a higher rate of innovation, and therefore a higher rate of creative destruction. Since the reward to success is decreasing in the rate of creative destruction, the private equilibrium will distort R&D labor away from the industries where it is more productive and towards industries where it is less productive.

Now suppose that under policy ψ_o, the private equilibrium allocates too little labor to R&D. Protection from future innovators, in the form of a patentability requirement and/or leading breadth may be able to stimulate R&D investment and the rate of innovation. If industries differ in their R&D capabilities, however, we must ask how such a policy will affect R&D distortions. The answer to this question is complicated in that it can depend on the bargaining solution. To illustrate the forces at work, we first ask what happens under the extreme bargaining solution where $\mathbf{s}^\alpha = (1, 0, \ldots, 0)$ for each α. Since the bargaining solution for $\alpha = 1$ is trivially $\mathbf{s}^1 = (1)$, this case includes all policies with a patentability requirement and no leading breadth. In addition, for policies with leading breadth we can interpret this assumption as a proxy for the rate of time preference ρ being very small relative to the industry arrival rate ϕ, in which case backloading becomes irrelevant. The following proposition shows that in this case stronger patent protection exacerbates R&D distortions.

Proposition 6 *(Stronger patent protection can exacerbate R&D distortions.) Suppose $h_1 > h_2$, so R&D labor is more productive in sector-1 industries. Suppose also that the bargaining solution is $\mathbf{s}^\alpha = (1, 0, \ldots, 0)$ for each α.*

If either $\hat{\phi}_1(\psi) > \hat{\phi}_1(\psi')$ or $\hat{\phi}_2(\psi) > \hat{\phi}_2(\psi')$ (i.e., if policy ψ increases the industry arrival rate in either sector relative to policy ψ'), then $\hat{N}_1(\psi)/\hat{N}_2(\psi) < \hat{N}_1(\psi')/\hat{N}_2(\psi') < [N_1/N_2]^*$.

Proposition 6 is also driven by the differential rates of creative destruction in the two sectors. By imposing stronger patent protection in order to induce firms to hire more R&D labor, policymakers increase the rate of creative destruction in all industries. But R&D labor is more productive in sector-1 industries, and as a result the increase in the rate of creative destruction is disproportionately large in sector-1 industries relative to sector-2 industries. Since the reward to success depends negatively on the rate of creative destruction, the private equilibrium will allocate the additional R&D labor with a bias towards sector-2 industries. Hence, stronger patent protection exacerbates the R&D distortion.

For policies with a patentability requirement (and no leading breadth), Proposition 6 implies that stimulating R&D investment will exacerbate R&D distortions. For policies with leading breadth, in contrast, Proposition 6 only tells part of the story, because the proposition relies on the extreme bargaining solution where the most recent innovator gets the entire bargaining surplus. The following example illustrates the importance of this assumption by demonstrating that the result can be overturned if we consider the alternative extreme:

Example Assume $L = 2.5$ and $\rho = 0.10$. Suppose that under policy ψ_o, firms choose innovation size $\gamma_o = 1.1$, and suppose further that $\lambda(\gamma_o) = 1$. Assume R&D production function $\phi = \lambda(\gamma)h^{1/2}n^{1/2}$, and suppose $h_1 = 2H_1 = 4$ and $h_2 = 2H_2 = 1$. Then we have $\phi_1 = 2n_1^{1/2}$ and $\phi_2 = n_2^{1/2}$, from which we can derive $N_1 = \frac{1}{2}n_1 = \frac{1}{8}\phi_1^2$, $N_2 = \frac{1}{2}n_2 = \frac{1}{2}\phi_2^2$, and $N = \frac{1}{8}\phi_1^2 + \frac{1}{2}\phi_2^2$.

Under policy ψ_o, the no-profit conditions are

$$(L - N)(\gamma - 1)\frac{1}{\rho + \phi_1} = \frac{\phi_1}{2} \quad \text{and} \quad (L - N)(\gamma - 1)\frac{1}{\rho + \phi_2} = 2\phi_2.$$

These no profit-conditions are satisfied by $\phi_1 = 0.645$ and $\phi_2 = 0.300$, which yields $\phi_1/\phi_2 = 2.15$.

Now consider policy $\psi' \equiv (P = \gamma_o, K = \gamma_o^2)$, and suppose the bargaining solution is $\mathbf{s}^2 = (0, 1)$. Then the no-profit conditions are

$$(L - N)(\gamma^2 - 1)\frac{\phi_1}{(\rho + \phi_1)^2} = \frac{\phi_1}{2} \quad \text{and} \quad (L - N)(\gamma^2 - 1)\frac{\phi_2}{(\rho + \phi_2)^2} = 2\phi_2.$$

These no profit-conditons are satisfied by $\phi_1 = 0.886$ and $\phi_2 = 0.393$, which yields $\phi_1/\phi_2 = 2.25$.

In this example, leading breadth stimulates the rate of innovation in both sectors, and moreover decreases the R&D distortion (a social planner would set $\phi_1/\phi_2 = 4$). Why does Proposition 6 fail to hold when we relax the assumption of $\mathbf{s}^\alpha = (1, 0, \ldots, 0)$? To answer this question, recall that increased leading breadth has two effects on the reward to success. On one hand, it increases available market profits. But on the other hand, increased leading breadth can cause payoffs to become backloaded. The backloading of payoffs is more costly the smaller is the industry arrival rate. Under the assumption $\mathbf{s}^\alpha = (1, 0, \ldots, 0)$, only the first effect of increased leading breadth is present, whereas once we relax this assumption, the backloading effect arises. It is the backloading effect that tends to counteract the result in Proposition 6. Since the industry arrival rate is larger in sector-1 industries than in sector-2 industries, the backloading of payoffs will be more costly in sector-2 industries. As a result, increased leading breadth has a bigger effect on sector-1 industries, which is exactly what is needed to reduce R&D distortions.

In sum, then, our analysis in this section suggests that when industries have asymmetric R&D capabilities, patent protection will tend to distort R&D investment away from those industries where it would be most productive. In addition, if policymakers attempt to stimulate aggregate R&D investment with protection from future innovators, they will alter this R&D distortion. A patentability requirement tends to increase R&D distortions. In contrast, leading breadth may increase or decrease R&D distortions, depending on the backloading of payoffs.

7. Discussion and Conclusion

In this paper, we have examined patent policy in a model of endogenous growth. For simplicity, we have couched our analysis within a relatively simple quality-ladder model along the lines of Grossman and Helpman (1991) and Aghion and Howitt (1992). Within this model, we are forced to make a number of simplifying assumptions to keep the analysis tractable. For instance, we use a very simple model of licensing, we assume there is no scope to have different patent policies apply to different industries, and we consider a world where market leaders do not have an incentive to conduct R&D. Moreover, the simple quality-ladder models of Aghion and Howitt and Grossman and Helpman may not be the most realistic of endogenous-growth models. Jones (1995) rejects R&D-based models of endogenous-growth because they exhibit scale effects that are inconsistent with time-series evidence, and a number of authors have suggested models that are more consistent with empirical evidence.

Even so, the basic intuitions we identify are not special to our model. In particular, the logic behind our results prevails as long as R&D takes place within product lines and as long as there is leapfrogging (market leaders are repeatedly replaced by innovators). While not all second- and third-generation endogenous-growth models feature these characteristics, we suspect that the reason for neglecting them is tractability, and not lack of realism. Because our policy conclusions survive in a third-generation endogenous-growth model that has these crucial characteristics, we believe our results are quite relevant for the endogenous-growth literature. We conclude by discussing some general lessons that the reader should take away from our analysis.

Perhaps the most basic lesson to take away from our analysis is that whenever R&D firms face a threat from future innovators, there may be a role for patents to provide protection against future innovators. In R&D-based growth models where successful R&D firms have their market profits eroded when other firms come along with new inventions, there can be a role for protection against future innovators in those models.

A second lesson is that, in addition to stimulating R&D investment, patent policy can also be useful for influencing the direction of firms' inventive activity. In our model, this took the form of imposing a patentability requirement to counteract firms' tendencies to pursue suboptimally small innovations. In a similar fashion, a patentability requirement and/or leading breadth could influence the characteristics of new products, or the types of cost reductions that firms pursue.

A third lesson is that any examination of government policy with regard to R&D must carefully assess the static efficiency implications of any policy proposal, and to do so one must consider the economy-wide implications of the policy. Our analysis suggests that any policy that will affect all industries equally or that will affect only a few industries may have small static efficiency implications, whereas any policy that has asymmetric effects across industries can have large static efficiency implications.

A fourth lesson is that the theoretical R&D literature may be missing an important issue when it assumes symmetric R&D capabilities across industries. Our model shows that patent policy can cause R&D distortions, and that the magnitude of those distortions depends on the specific patent policy. However, this lesson does not apply only to analyses of patent policy. Any analysis of government policy for R&D should ask what are the implications for R&D distortions.

Perhaps our most important contribution is the merging of the patent-design literature and the endogenous-growth literature. The patent-design literature focusses on specific policy instruments, with attention paid to institutional detail. The endogenous-growth literature builds careful models of R&D at the economy-wide level, with attention paid to empirical calibration. By merging these literatures, we hope this paper will help improve the study of government policy with regard to R&D.

Appendix 1: Equilibrium with Leading Breadth

Equilibrium Conditions in the Basic Model (Section 3)

Combining Lemmas 2 and 3, under any patent policy $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = P^\alpha$, the markup in high-technology industries will be $\mu = P^\alpha$ and the reward to success will be $V = \pi \cdot B(\phi, \alpha)$ where $B(\phi, \alpha) \equiv \Sigma_{i=1}^{\alpha} s_i^\alpha \phi^{i-1}/(\rho + \phi)^i$ and $\pi = (L - N)(\mu - 1)/((1 - \bar{\omega})\mu + \bar{\omega})$. Any ϕ that satisfies the no-profit condition $\lambda V = w$ is a candidate for an ''equilibrium''. Given $N = \bar{\omega} n = (\bar{\omega}/\lambda)\phi$, define $\pi(\phi) = (L - (\bar{\omega}/\lambda)\phi)(\mu - 1)/((1 - \bar{\omega})\mu + \bar{\omega})$. We can then write the no-profit condition as

$$B(\phi, \alpha) = \frac{w}{\lambda \pi(\phi)}.$$

Closer inspection of this equation reveals that ϕ is the only endogenous variable, as should be the case. Note that $\pi(\phi)$ is decreasing in ϕ, and that $\lim_{\phi \to (\lambda/\bar{\omega})L} \pi(\phi) = 0$ (where $\phi = (\lambda/\bar{\omega})L$ implies $N = L$). This implies $w/(\lambda\pi(\phi))$ is increasing in ϕ and $\lim_{\phi \to (\lambda/\bar{\omega})L} w/(\lambda\pi(\phi)) = \infty$.

What is $B(\phi, \alpha)$? If $\alpha = 1$ then $B(\phi, \alpha) = 1/(\rho + \phi)$, which is everywhere decreasing in ϕ. As a result, for any policy without leading breadth (i.e., $\alpha = 1$) there is a unique ϕ that satisfies the no-profit condition, as illustrated in Figure 1.

If $\alpha > 1$ then $B(\phi, \alpha)$ can sometimes be everywhere decreasing in ϕ, in which case a unique ϕ satisfies the no-profit condition. For example, $\mathbf{s}^\alpha = (1/\alpha, 1/\alpha, \ldots, 1/\alpha)$ implies $B(\phi, \alpha) = \Sigma_{i=1}^{\alpha} (1/\alpha)\phi^{i-1}/(\rho + \phi)^i = (1/(\alpha\rho))[1 - (\phi/(\rho + \phi))^\alpha]$, which is everywhere decreasing in ϕ. But $\alpha > 1$ can also yield $B(\phi, \alpha)$ not everywhere decreasing in ϕ. For example, if the bargaining solution $\mathbf{s}^2 = (0, 1)$ then $B(\phi, \alpha) = \phi/(\rho + \phi)^2$, which is initially increasing in ϕ and then decreasing in ϕ. In this case, multiple ϕ can satisfy the no-profit condition, as illustrated in Figure 2. When multiple ϕ satisfy the no-profit condition, we define the largest such ϕ to be the equilibrium.[25]

For all α, $B(\phi, \alpha)$ is continuous in ϕ, and also $\lim_{\phi \to (\lambda/\bar{\omega})L} B(\phi, \alpha)$ is finite (to see the latter note that $B(\phi, \alpha) \leq 1/(\rho + \phi)$ for any α). Two results follow. First, our equilibrium definition is well-defined—that is, there must be some $\hat{\phi} < (\lambda/\bar{\omega})L$ such that $B(\hat{\phi}, \alpha) = w/(\lambda\pi(\hat{\phi}))$ and $B(\phi, \alpha) < w/(\lambda\pi(\phi))$ for all $\phi > \hat{\phi}$ (unless $B(\phi, \alpha) < w/(\lambda\pi(\phi))$ for all $\phi < (\lambda/\bar{\omega})L$, in which case there is no steady-state R&D).

25 In Figure 2, $\hat{\phi}_1$ is unstable in the sense that ϕ slightly larger implies $\lambda V > w$, which would lead all firms to increase R&D spending further until $\hat{\phi}_2$ is reached. The largest ϕ that satisfies the no-profit condition is always stable. For $\alpha = 2$, the largest ϕ that satisfies the no-profit condition is in fact the only stable solution; for $\alpha \geq 3$, there could be additional (intermediate) stable solutions.

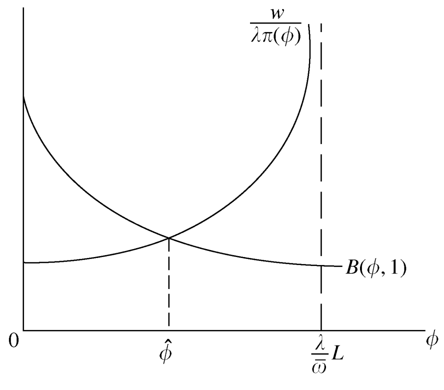

Figure 1. Equilibrium when $\alpha = 1$.

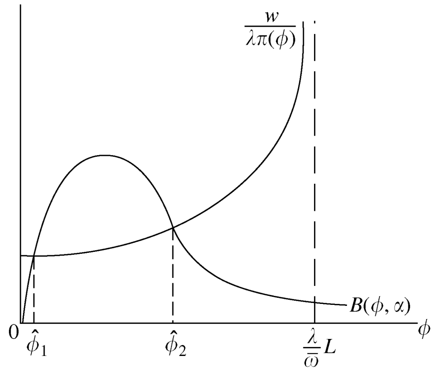

Figure 2. Possible equilibria when $\alpha > 1$.

Second, if $B(\phi', \alpha) > 1/(\lambda\pi(\phi'))$, then $\hat{\phi}(\psi) > \phi'$. In words, if for some ϕ' the reward to success is larger than the wage, then the equilibrium industry arrival rate is larger than ϕ'. This second result will be useful in proving Propositions 2 and 3.

Equilibrium Conditions in the Asymmetric-R&D Model (Section 6)

Under any patent policy $\psi = (P, K)$ with $P \geq \gamma_o$ and $K = P^\alpha$, all firms target innovation size P, and the markup in all industries will be $\mu = P^\alpha$. The reward to success will be $V = \bar{\pi} \cdot B(\phi, \alpha)$ where $B(\phi, \alpha) \equiv \Sigma_{i=1}^{\alpha} s_i^\alpha \phi^{i-1}/(\rho + \phi)^i$ and $\bar{\pi} = (L - N)(\mu - 1)$. Any combination (n_1, n_2) that satisfies the no-profit conditions $\beta\lambda(h_1/n_1)^{1-\beta} V_1 = w$ and $\beta\lambda(h_2/n_2)^{1-\beta} V_2 = w$ is a candidate for an ''equilibrium''.

It is convenient to reframe the problem as finding a combination (N, ϕ_1, ϕ_2) that must satisfy the condition

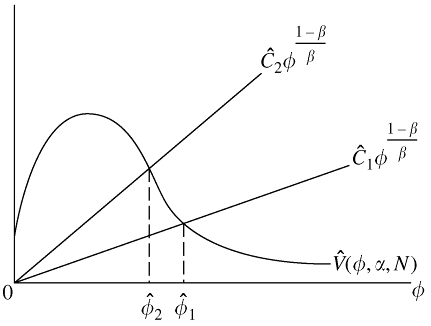

Figure 3. Equilibrium in asymmetric-R&D model.

$$N = \frac{1}{2}\lambda^{-1/\beta}h_1^{(1-\beta)/-\beta}\phi_1^{1/\beta} + \frac{1}{2}\lambda^{-1/\beta}h_2^{(1-\beta)/-\beta}\phi_2^{1/\beta}. \tag{A.1}$$

Define $\hat{V}(\phi, \alpha, N) \equiv (L - N)(P^{\alpha} - 1)B(\phi, \alpha)$, and note that this function is the same for both sectors, so any differences in the reward to success across sectors are driven by different industry arrival rates. Define $\hat{C}_j \equiv w[\beta\lambda^{1/\beta}h_j^{(1-\beta)/\beta}]^{-1}$, and note that $h_1 > h_2$ implies $\hat{C}_1 < \hat{C}_2$. We can then rewrite the no-profit conditions as

$$\hat{V}(\phi_1, \alpha, N) = \hat{C}_1\phi_1^{(1-\beta)/\beta} \quad \text{and} \quad \hat{V}(\phi_2, \alpha, N) = \hat{C}_2\phi_2^{(1-\beta)/\beta}.$$

These no-profit conditions are graphed in Figure 3. As in the basic model, $B(\phi, \alpha)$ can be non-monotonic in ϕ and therefore $\hat{V}(\phi, \alpha, N)$ can be non-monotonic in ϕ. As a result, for any fixed N, multiple ϕ can satisfy the no-profit condition for sector j. We again define the largest such ϕ to be the equilibrium—that is, for any fixed N,

$$\hat{\phi}_j \equiv \begin{cases} \max\{\phi \mid V_j = \hat{C}_j\phi^{(1-\beta)/\beta}\} & \text{if the set } \{\phi \mid V_j = \hat{C}_j\phi^{(1-\beta)/\beta}\} \text{ is non-empty} \\ 0 & \text{if the set } \{\phi \mid V_j = \hat{C}_j\phi^{(1-\beta)/\beta}\} \text{ is empty.} \end{cases}$$

Since $\partial\hat{V}/\partial N < 0$ and $\hat{C}_1\phi^{(1-\beta)/\beta}$ and $\hat{C}_2\phi^{(1-\beta)/\beta}$ are independent of N, the no-profit conditions imply that (all else equal) ϕ_1 and ϕ_2 are decreasing in N. We can therefore conclude that for any patent policy ψ a unique combination $(\hat{N}(\psi), \phi_1(\psi), \phi_2(\psi))$ satisfies the no-profit conditions and equation (A.1). Moreover, since $\hat{C}_1 < \hat{C}_2$ implies $\hat{V}(\hat{\phi}_2(\psi), \alpha, \hat{N}(\psi)) = \hat{C}_2(\hat{\phi}_2(\psi))^{(1-\beta)/\beta} > \hat{C}_1(\hat{\phi}_2(\psi))^{(1-\beta)/\beta}$, we can conclude that $h_1 > h_2$ implies $\hat{\phi}_1(\psi) > \hat{\phi}_2(\psi)$.

Appendix 2: Proofs

Proof of Proposition 1. (i) Define $\tilde{N} \equiv \bar{\omega}[((P-1)/P)L - (((1-\bar{\omega})P + \bar{\omega})/P)(\rho/\lambda(P))]$, and then Lemma 2 implies $\hat{N}(\psi) = \max\{\tilde{N}, 0\}$ and $\hat{\phi}(\psi) = \lambda(P)\hat{N}(\psi) = \max\{\lambda(P)\tilde{N}, 0\}$. In addition, whenever $\tilde{N} \geq 0$, $\hat{N}(\psi)$ is increasing in P if and only if $d\tilde{N}/dP > 0$, and $\hat{\phi}(\psi)$ is increasing in P if and only if $d[\lambda(P)\tilde{N}]/dP > 0$. Differentiating:

$$\frac{d\tilde{N}}{dP} = \bar{\omega}\left[\frac{1}{P^2}L + \frac{\bar{\omega}}{P^2}\frac{\rho}{\lambda} - \frac{(1-\bar{\omega})P + \bar{\omega}}{P}\frac{\rho}{\lambda}\left(\frac{-d\lambda/dP}{\lambda}\right)\right]$$

and

$$\frac{d[\lambda(P)\tilde{N}]}{dP} = \bar{\omega}\left[\frac{d\lambda}{dP}\left(\frac{P-1}{P}\right)L + \frac{\lambda}{P^2}L + \frac{\bar{\omega}}{P^2}\rho\right].$$

$N_o > 0$ implies $L \geq (((1-\bar{\omega})\gamma_o + \bar{\omega})/(\gamma_o - 1))(\rho/\lambda)$, and at $P = \gamma_o$ we have $-d\lambda/dP = \lambda/(\gamma_o(\gamma_o - 1))$. It is then straightforward to show $d\tilde{N}/dP|_{\gamma_o} > 0$ and $d[\lambda(P)\tilde{N}]/dP|_{\gamma_o} > 0$, from which it follows that there exists $P' > \gamma_o$ and $P'' > \gamma_o$ such that $\hat{N}(\psi)$ is increasing in P for all $P \in [\gamma_o, P')$ and $\hat{\phi}(\psi)$ is increasing in P for all $P \in [\gamma_o, P'')$. That $P' > P''$ follows from $d[\lambda(P)\tilde{N}]/dP = \lambda d\tilde{N}/dP + d\lambda/dP < 0$ at $P = P'$.

(ii) This result follows from $d\tilde{N}/dP|_{\gamma_o} > 0$ and $d[\lambda(P)\tilde{N}]/dP|_{\gamma_o} > 0$ combined with the fact that $\tilde{N}$ is continuous and differentiable, because then either $\tilde{N} \leq 0$ for all $P \geq \gamma_o$ or there exists $\underline{P} \geq \gamma_o$ such that $P = \underline{P}$ implies $\tilde{N} = 0$ and $d\tilde{N}/dP > 0$. ■

Proof of Lemma 3. Suppose for now that all R&D firms choose innovation size $\gamma = P$ and therefore there are α firms in each licensing agreement; we will later prove that $\hat{\gamma}(\psi) = P$. If there are α firms in each licensing agreement, $\mu = P^\alpha$ in all high-technology industries. Given the markup μ, market profits in all high-technology industries are given by π (see equation (8)). Given $\mathbf{s}^\alpha$ and π, a successful firm receives flow payoff $s_1^\alpha \pi$ until the first subsequent innovation, then flow payoff $s_2^\alpha \pi$ until the second subsequent innovation, and so on. Since the reward to success V is the discounted value of this payoff stream, if the industry arrival rate is ϕ then $V = \pi \Sigma_{i=1}^{\alpha} s_i^\alpha \phi^{i-1}/(\rho + \phi)^i = \pi \cdot B(\phi, \alpha)$ (see the calculations in footnote 7).

It remains to prove that $\hat{\gamma}(\psi) = P$. While a firm clearly would not choose innovation size $\gamma < P$, there are two reasons that a firm might choose innovation size $\gamma > P$: to affect marginal profits, and to avoid the leading breadth of some prior patent and thereby change the number of firms in each licensing agreement. Consider the latter. This incentive is present only if $\alpha \geq 2$, since otherwise innovation size $\gamma = P$ avoids the leading breadth of all prior patents. Moreover, if all other firms choose $\gamma = P$, then to avoid the leading breadth of some prior patent a firm must choose $\gamma \geq P^2$. But $d^2\lambda/d\gamma^2 \leq 0$ implies $(\lambda(\gamma_o^2) - \lambda(\gamma_o))/(\gamma_o^2 - \gamma_o) \leq d\lambda/d\gamma\,|_{\gamma=\gamma_o}$, and substituting equation (6) this inequality implies $\lambda(\gamma_o^2) = 0$. Hence, for any $P \geq \gamma_o$, innovations of size $\gamma \geq P^2$ have $\lambda(\gamma) = 0$ (they are not feasible), and hence a firm would never choose $\gamma > P$ so as to avoid the leading breadth of some prior patent.

Now consider whether a firm would choose $\gamma > P$ to affect marginal profits. Suppose all other firms choose innovation size $\gamma' \geq 1$. If a firm has an innovation of size γ, then the industry markup while the firm is part of the bargaining group will be $\mu = \Gamma\gamma$, where $\Gamma = (\gamma')^{\alpha-1} \geq 1$. The R&D firm therefore has instantaneous payoff

$$-wn + \lambda(\gamma)nLY\left(\frac{\Gamma\gamma - 1}{\Gamma\gamma}\right)B(\phi, \alpha),$$

where the firm takes L, Y, w, Γ, and $B(\phi, \alpha)$ as given and chooses γ and n subject to $\gamma \geq P$.

Since $P \geq \gamma_o$, $\lambda(\gamma)((\gamma - 1)/\gamma)$ is decreasing at $\gamma = P$. It is then straightforward to show that $\Gamma \geq 1$ implies $\lambda(\gamma)(\Gamma\gamma - 1)/(\Gamma\gamma)$ is also decreasing at $\gamma = P$, and hence innovation size $\gamma = P$ is optimal (given the constraint $\gamma \geq P$). The result follows. ■

Proof of Proposition 2. Let α' and α'' be such that $K = P^{\alpha'}$ and $K' = P^{\alpha''}$. Define $\pi(\phi, \alpha) \equiv (L - (\bar{\omega}/\lambda)\phi)\ (P^\alpha - 1)/((1 - \bar{\omega})P^\alpha + \bar{\omega})$ and let $\bar{\phi} = \hat{\phi}(\psi)$, which implies $\lambda\pi(\bar{\phi}, \alpha') \cdot B(\bar{\phi}, \alpha') = w$. Using the logic from Appendix A, if $\lambda\pi(\bar{\phi}, \alpha'') \cdot B(\bar{\phi}, \alpha'') \geq \lambda\pi(\bar{\phi}, \alpha') \cdot B(\bar{\phi}, \alpha') = w$, then $\hat{\phi}(\psi') > \hat{\phi}(\psi)$.

i. If $\alpha' = 1$ and $\alpha'' = 2$, then $\hat{\phi}(\psi') > \hat{\phi}(\psi)$ if

$$\left(L - \frac{\bar{\omega}}{\lambda}\bar{\phi}\right)\frac{(P^2 - 1)}{(1 - \bar{\omega})P^2 + \bar{\omega}}\frac{s\rho + \bar{\phi}}{(\rho + \bar{\phi})^2} \geq \left(L - \frac{\bar{\omega}}{\lambda}\bar{\phi}\right)\frac{(P - 1)}{(1 - \bar{\omega})P + \bar{\omega}}\frac{1}{(\rho + \bar{\phi})}.$$

We can rewrite this inequality as $s \geq \Gamma - (\bar{\phi}/\rho)[1 - \Gamma] \equiv \bar{s}$, where $\Gamma \equiv (1/(P + 1))((1 - \bar{\omega})P^2 + \bar{\omega})/((1 - \bar{\omega})P + \bar{\omega}) < 1$. $\bar{s}$ is maximized at $\bar{\omega} = 0$, where $\bar{s} = P/(P + 1) - (\bar{\phi}/\rho)(1/(P + 1)) \leq P/(P + 1)$.

ii. Suppose $\alpha' = \alpha$ and $\alpha'' = \alpha + 1$. Since $\phi^{\alpha-1}/(\rho + \phi)^\alpha \leq B(\phi, \alpha) \leq 1/(\rho + \phi)$ for all α and ϕ, we must have $\hat{\phi}(\psi') > \hat{\phi}(\psi)$ if

$$\left(L - \frac{\bar{\omega}}{\lambda}\bar{\phi}\right)\frac{(P^{\alpha+1} - 1)}{(1 - \bar{\omega})P^{\alpha+1} + \bar{\omega}}\frac{\bar{\phi}^\alpha}{(\rho + \bar{\phi})^{\alpha+1}} \geq \left(L - \frac{\bar{\omega}}{\lambda}\bar{\phi}\right)\frac{(P^\alpha - 1)}{(1 - \bar{\omega})P^\alpha + \bar{\omega}}\frac{1}{(\rho + \bar{\phi})}.$$

We can rewrite this inequality as

$$\rho \leq \bar{\phi}\left[\left(\frac{(P^{\alpha+1} - 1)}{(P^\alpha - 1)}\frac{(1 - \bar{\omega})P^\alpha + \bar{\omega}}{(1 - \bar{\omega})P^{\alpha+1} + \bar{\omega}}\right)^{1/\alpha} - 1\right] \equiv \bar{\rho}.$$

It is straightforward to show $\bar{\rho} > 0$, and the result follows. ■

Proof of Proposition 3. Appendix A shows that the equilibrium industry arrival rate $\hat{\phi}$ must satisfy the no-profit condition $B(\hat{\phi}, \alpha) = w/(\lambda\pi(\hat{\phi}))$, and moreover that if $B(\phi', \alpha) > w/(\lambda\pi(\phi'))$ then $\hat{\phi} > \phi'$. Since $B(\phi, \alpha)$ and λ are independent of $\bar{\omega}$, the result follows if $\pi(\phi)$ is increasing in $\bar{\omega}$. From Appendix A, $\pi(\phi) = (L - (\bar{\omega}/\lambda)\phi)$ $(\mu - 1)/((1 - \bar{\omega})\mu + \bar{\omega})$, and differentiating yields $\partial\pi(\phi)/\partial\bar{\omega} = (\mu - 1)/[(1 - \bar{\omega})\mu + \bar{\omega}]^2\ [(\mu - 1)L - \mu\phi/\lambda]$, so $\partial\pi(\phi)/\partial\bar{\omega} > 0$ if and only if $\phi < ((\mu - 1)/\mu)\lambda L$. Since $B(\phi,\ \alpha) \leq 1/(\rho + \phi)$ for all ϕ and α, we must have $\hat{\phi} < \bar{\phi}$ where $\bar{\phi}$ is defined by $1/(\rho + \bar{\phi}) = 1/(\lambda\pi(\bar{\phi}))$. Solving this latter equation for $\bar{\phi}$ yields $\bar{\phi} = ((\mu - 1)/\mu)\lambda L - (((1 - \bar{\omega})\mu + \bar{\omega})/\mu)\rho < ((\mu - 1)/\mu)\lambda L$, and the result follows. ■

Proof of Proposition 4. Under any policy $\psi = (P,\ K)$ with $P \geq \gamma_o$ and $K = P^\alpha$, the markup $\mu = P^\alpha$ regardless of $\bar{\omega}$. That there is no static inefficiency for $\bar{\omega} = 0$ and $\bar{\omega} = 1$ follows directly from equation (11). More generally, the static inefficiency is increasing in

$\bar{\omega}$ if and only if $\partial D/\partial\bar{\omega} > 0$, and the static inefficiency is decreasing in $\bar{\omega}$ if and only if $\partial D/\partial\bar{\omega} < 0$. Differentiating equation (11) yields

$$\frac{\partial D}{\partial\bar{\omega}} = \frac{\mu\ln\mu - (\mu - 1)(1 + \bar{\omega}\ln\mu)}{(1 - \bar{\omega})\mu + \bar{\omega}}.$$

$\partial D/\partial\bar{\omega} > 0$ if and only if $\bar{\omega} < \mu/(\mu - 1) - 1/\ln\mu$, and $\partial D/\partial\bar{\omega} < 0$ if and only if $\bar{\omega} > \mu/(\mu - 1) - 1/\ln\mu$. For any $\mu > 1$, we have $0 < \mu/(\mu - 1) - 1/\ln\mu < 1$, and the result follows. ■

Proof of Proposition 5. Combining equations (16) and (17), and using $N_j = \frac{1}{2}n_j$ and $H_j = \frac{1}{2}h_j$, in the private equilibrium we must have under any policy ψ

$$\frac{V_1}{V_2} = \left(\frac{H_2}{H_1}\right)^{1-\beta}\left(\frac{N_1}{N_2}\right)^{1-\beta} \quad \text{or} \quad \frac{N_1}{N_2} = \frac{H_1}{H_2}\left(\frac{V_1}{V_2}\right)^{1/(1-\beta)}.$$

Under policy ψ_o, equation (15) implies that $V_1/V_2 = (\rho + \phi_2)/(\rho + \phi_1)$. In Appendix A, we show that $h_1 > h_2$ implies $\phi_1 > \phi_2$, and so $V_1/V_2 < 1$. Since $[N_1/N_2]^* = H_1/H_2$, $V_1 < V_2$ implies $\hat{N}_1(\psi_o)/\hat{N}_2(\psi_o) < [N_1/N_2]^*$. ■

Proof of Proposition 6. $s^\alpha = (1, 0, \ldots, 0)$ implies that $V_1/V_2 = (\rho + \phi_2)/(\rho + \phi_1)$. As in the proof of Proposition 5, this implies that $\hat{N}_1(\psi)/\hat{N}_2(\psi) < [N_1/N_2]^*$ for any ψ. Also note that $\phi_j = \lambda h_j^{1-\beta} n_j^\beta$ implies that $z(\psi) \equiv \hat{\phi}_1(\psi)/\hat{\phi}_2(\psi) = (H_1/H_2)^{1-\beta}(\hat{N}_1(\psi)/\hat{N}_2(\psi))^\beta > 1$ (where the inequality follows from $h_1 > h_2$ implying $\hat{N}_1(\psi) > \hat{N}_2(\psi)$).

Now suppose that either $\hat{\phi}_1(\psi) > \hat{\phi}_1(\psi')$ or $\hat{\phi}_2(\psi) > \hat{\phi}_2(\psi')$. The claim is that this implies $\hat{N}_1(\psi)/\hat{N}_2(\psi) < \hat{N}_1(\psi')/\hat{N}_2(\psi')$. Suppose otherwise. $\hat{N}_1(\psi)/\hat{N}_2(\psi) \geq \hat{N}_1(\psi')/\hat{N}_2(\psi')$ implies $z(\psi) \geq z(\psi')$. Because we can rewrite

$$\frac{V_1}{V_2} = \frac{\rho + \phi_2}{\rho + \phi_1} = \frac{\rho + \frac{1}{z}\phi_1}{\rho + \phi_1} = \frac{\rho + \phi_2}{\rho + z\phi_2}, z(\psi) \geq z(\psi')$$

and either $\hat{\phi}_1(\psi) > \hat{\phi}_1(\psi')$ or $\hat{\phi}_2(\psi) > \hat{\phi}_2(\psi')$ implies V_1/V_2 must be smaller under policy ψ. But since the proof of Proposition 5 implies that N_1/N_2 is strictly increasing in V_1/V_2, this implies $\hat{N}_1(\psi)/\hat{N}_2(\psi) < \hat{N}_1(\psi')/\hat{N}_2(\psi')$, a contradiction. The result follows. ■

Appendix 3: A Non-Scale Growth Model

In this appendix, we sketch an endogenous-growth model without scale effects—similar to Howitt (1999) and Aghion and Howitt (1998, chapter 12)—in which the patent policies described in Section 3 can be effective at stimulating R&D.

Modified Model

Our model here mirrors our basic framework with four modifications. First, to simplify the analysis, we assume there is no noninnovative sector. Second, we assume that the labor force is growing at exogenous rate g ($\dot{L}/L = g$). Third, and most importantly, we permit the number of varieties to grow over time. Formally, the set of varieties at time t is $[0, B(t)]$. Preferences take the same basic form, where the instantaneous utility function is

$$\ln u(t) = \int_0^{B(t)} \ln[q_\omega(t)\ x_\omega(t)]\, d\omega.$$[26]

Fourth, we endogenize the growth rate of new varieties by introducing horizontal R&D. We assume that, when a new variety is invented, a markup $\bar{\mu}$ can be charged until someone comes along with a new improved version of that variety. The magnitude of $\bar{\mu}$ depends on the strength of lagging breadth for patents on new varieties. We take $\bar{\mu}$ to be exogenous, or, more to the point, to be independent of protection against future innovators (which is consistent with our earlier assumptions). Horizontal R&D is carried out by individual workers. We follow Howitt (1999) in assuming aggregate diminishing returns to horizontal R&D due to workers' unequal abilities. Specifically, each worker is endowed with a horizontal R&D ability h_i, which means that the worker's horizontal R&D production function is $\dot{B}_i = h_i$. The cross-sectional distribution of h_i is constant over time with cumulative density $F(h)$; we assume that F is continuous and increasing on domain $[0, \infty)$, which guarantees an interior solution for horizontal R&D. Workers are identical in the other tasks, manufacturing and vertical R&D. Hence, if at time t the reward from a horizontal invention is $V_h(t)$ (which, as discussed below, will be endogenously determined), then at time t type h_i will devote his labor supply to horizontal R&D if and only if $h_i V_h(t) \geq w$ (because $h_i V_h(t)$ is the expected benefit and w is the opportunity cost from not entering the normal labor market). It follows that all workers with $h_i \geq h^* \equiv w/V_h(t)$ will engage in horizontal R&D, whereas all workers with $h_i < h^*$ will work either in manufacturing or in vertical R&D. As a result, the number of horizontal R&D workers at time t will be $N_h = L(1 - F(h^*))$.[27]

Vertical R&D invents new improved versions of existing products, and proceeds exactly

26 Under this specification, which is analogous to that in Howitt (1999), both vertical and horizontal innovations improve utility. Alternatively, we might follow Aghion and Howitt (1998, chapter 12) and normalize the utility function such that horizontal innovations are neutral—for example, we could assume that

$$\ln u(t) = \frac{1}{B(t)} \int_0^{B(t)} \ln[q_\omega(t) x_\omega(t)]\, d\omega.$$

Any such normalization would not affect the market outcome, but would have implications for any normative analysis of that market outcome.

27 Our formulation creates income inequality as more productive horizontal R&D workers get a higher reward. However, this has no implications for the equilibrium because consumers have homothetic preferences. Only aggregate income, and not its distribution, matter for the demand curve of (horizontal and vertical) innovators.

as in our basic model. To clarify notation, we use $N_v(t)$ and $n_v(t)$ to denote economy-wide vertical R&D and industry-level vertical R&D, respectively (and so $N_v(t) = B(t)n_v(t)$).

The Private Equilibrium under Policy ψ_o

We restrict attention to balanced-growth steady states in which $\dot{B}/B = \dot{L}/L \equiv g$. In such steady states, the amount of manufacturing labor per product n_m and the amount of vertical R&D per product n_v will be constant over time. Economy-wide manufacturing labor $N_m(t) = B(t)n_m$, economy-wide vertical R&D $N_v(t) = B(t)n_v$, and economy-wide horizontal R&D $N_h(t)$ will all grow at rate g.

The analysis is much as in the basic model. As in the basic model, steady-state per-person expenditures Y are constant over time. Because each person allocates expenditures evenly across existing industries, the per-person demand for variety ω is $Y/(Bp_\omega)$, and therefore the demand function for variety ω is $LY/(Bp_\omega)$. Since L/B is constant over time, this demand function is constant over time. If the market leader for industry ω can charge markup μ_ω, it will earn profit $\pi_\omega = (LY/B)((\mu_\omega - 1)/\mu_\omega)$. The markup may differ for first-generation products (of that variety) vs. later-generation products. As discussed above, we assume the markup for first-generation products is $\bar{\mu}$, and so flow profit following a horizontal innovation is $\pi_h = (LY/B)((\bar{\mu} - 1)/\bar{\mu})$. Under policy ψ_o, if vertical R&D firms choose innovation size γ, then profits are $\pi_v = (LY/B)((\gamma - 1)/\gamma)$.

Let ϕ_v denote the per-industry arrival rate for vertical innovations (which is constant over time given n_v is constant over time). For each vertical R&D firm, as a function of its innovation size γ and R&D spending n, its instantaneous payoff will be

$$-wn + \lambda(\gamma)n\frac{\pi_v}{\rho + \phi_v} = -wn + \lambda(\gamma)n\left(\frac{LY}{B}\right)\left(\frac{\gamma - 1}{\gamma}\right)\Big/(\rho + \phi_v).$$

Each firm will choose innovation size γ_o exactly as in the basic model, and the free-entry condition is $\lambda(\gamma_o)\pi_v/(\rho + \lambda(\gamma_o)n_v) = 1$ (we again use the wage as the numeraire).

In the steady state, a proportion z of industries will have first-generation products, while a proportion $1 - z$ of industries will have later-generation products, where $z = g/(g + \phi_v)$. Each industry with a first-generation product will have markup $\bar{\mu}$, and therefore will hire manufacturing labor equal to $LY/(B\bar{\mu})$. Similarly, each industry with a later-generation product will have markup γ_o, and therefore will hire manufacturing labor equal to $LY/(B\gamma_o)$. Hence, economy-wide manufacturing labor will be $N_m = LY(z/\bar{\mu} + (1 - z)/\gamma_o)$. For the illustrative purposes of this appendix, we make the simplifying assumption that the proportion of industries z that currently has a first-generation product is approximately equal to zero—that is, we analyze a situation where the population growth rate is small relative to rate of quality improvements.[28] When $z \approx 0$, manufacturing labor per product is $n_m = LY/(B\gamma_o)$, and economy-wide manufacturing labor is

28 In fact, our analysis goes through when we properly account for the endogeneity of z—that is, our qualitative results below on the effects of a patentability requirement and on the effects of leading breadth are unchanged. Because policy changes affect z, however, there would be additional feedback

$N_m = Bn_m = LY/\gamma_o$. Note that in the steady state, n_m will be constant while N_m will grow at rate g.

The resource constraint requires $L = N_m + N_v + N_h = LY/\gamma_o + N_v + N_h$ at each date, or $Y = \gamma_o(L - N_h - N_v)/L$ (because L, N_v, and N_h all grow at rate g, Y is indeed constant). It follows that, along the balanced-growth path, profits for later-generation products are given by $\pi_v = (L - N_v - N_h)(\gamma_o - 1)/B$, which is constant over time. Plugging this into the free-entry condition and solving for vertical R&D employment per industry yields (for the interior case)

$$n_v = \frac{\gamma_o - 1}{\gamma_o}\left(\frac{L - N_h}{B}\right) - \frac{1}{\gamma_o}\frac{\rho}{\lambda(\gamma_o)}.$$

This is identical to our basic model with $\bar{\omega} = 1$, except that now N_h and B show up (the basic model assumed $N_h = 0$ and $B = 1$ at all dates).

As discussed above, the number of horizontal R&D workers at time t will be $N_h = L(1 - F(h^*))$, where $h^* \equiv w/V_h(t)$ and $V_h(t)$ is the reward from a horizontal invention at time t. Because a horizontal invention enables a firm to earn profit π_h until the first improvement, which occurs with arrival rate $\lambda(\gamma_o)n_v$, the reward to horizontal R&D is constant and equal to $V_h = \pi_h/(\rho + \lambda(\gamma_o)n_v)$. Hence, h^* satisfies

$$w = h^*\left[\frac{(LY/B)((\bar{\mu} - 1)/\bar{\mu})}{\rho + \lambda(\gamma_o)n_v}\right].$$

Recalling that the free-entry condition for vertical R&D is

$$w = \lambda(\gamma_o)\left[\frac{(LY/B)((\gamma_o - 1)/\gamma_o)}{\rho + \lambda(\gamma_o)n_v}\right],$$

it follows that

$$h^* = \lambda(\gamma_o)\left(\frac{\gamma_o - 1}{\gamma_o}\right)\bigg/\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right).$$

Finally, we solve for the steady-state value of L/B. The increase in the number of products is $\dot{B} = L\int_{h^*}^{\infty} h\,dF(h)$. Since in the steady state B must grow at rate g, we have

$$g = \frac{\dot{B}}{B} = \frac{L}{B}\int_{h^*}^{\infty} h\,dF(h) \quad \text{or} \quad \frac{L}{B} = g\bigg/\int_{h^*}^{\infty} h\,dF(h).$$

---

effects that do not show up in our analysis below (while these feedback effects can enhance or mitigate our conclusions, they do not change the basic qualitative conclusions). We have chosen to abstract away such effects because the equations and analysis become significantly more cumbersome.

We can now summarize the steady-state outcome under policy ψ_o (for the case when $n_v > 0$):

$$h^* = \lambda(\gamma_o)\left(\frac{\gamma_o - 1}{\gamma_o}\right)\bigg/\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right)$$
$$N_h/L = 1 - F(h^*)$$
$$n_v = \frac{\gamma_o - 1}{\gamma_o}\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\,dF(h)}\right) - \frac{1}{\gamma_o}\frac{\rho}{\lambda(\gamma_o)}$$
$$n_m = \frac{1}{\gamma_o}\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\,dF(h)}\right) + \frac{1}{\gamma_o}\frac{\rho}{\lambda(\gamma_o)}$$

The Effects of a Patentability Requirement

Much as in the basic model, the outcome under any policy with $P \geq \gamma_o$ and $K = 1$ is the same as under policy ψ_o except that P replaces γ_o. Hence, the steady-state outcome under any such policy will be (for the case when $n_v > 0$):

$$h^* = \lambda(P)\left(\frac{P - 1}{P}\right)\bigg/\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right)$$
$$N_h/L = 1 - F(h^*)$$
$$n_v = \frac{P - 1}{P}\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\,dF(h)}\right) - \frac{1}{P}\frac{\rho}{\lambda(P)}$$
$$n_m = \frac{1}{P}\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\,dF(h)}\right) + \frac{1}{P}\frac{\rho}{\lambda(P)}.$$

It is straightforward to show that imposing a patentability requirement can increase vertical R&D spending—formally, there exists $P' > \gamma_o$ such that n_v is increasing in P for all $P \in [\gamma_o, P')$. Intuitively, the logic behind Proposition 1 is unchanged, except that there is one additional effect. An increase in the patentability requirement reduces the reward to vertical R&D relative to the reward to horizontal R&D, and therefore there is an increase in horizontal R&D (decrease in h^*). While this increase in horizontal R&D reduces the impact of a patentability requirement on vertical R&D, this effect is second-order at $P = \gamma_o$ and therefore does not overturn our conclusion.[29]

29 Formally,

$$\frac{dn_v}{dP} = \frac{\partial n_v}{\partial P} + \frac{\partial n_v}{\partial h^*}\frac{dh^*}{dP}.$$

One can show that $\partial n_v/\partial P|_{P=\gamma_o} > 0$ using an identical technique to that in the proof of Proposition 1. One can also show $dh^*/\partial P|_{P=\gamma_o} = 0$, and therefore $dn_v/dP|_{P=\gamma_o} > 0$.

The Effects of Leading Breadth

The introduction of leading breadth—policies with $P \geq \gamma_o$ and $K = P^{\alpha}$—into our modified model creates significant further complications. First, instead of the markup differing only for first-generation products vs. later-generation products, here the markups for the first α generations differ from each other and from that for later generations—for example, when $\alpha = 2$, the markup for first-generation products is $\bar{\mu}$, the markup for second-generation products is $\bar{\mu}P$, and the markup for all later-generation products is P^2. Second, we must account for how profits for early-generation products are divided among patent holders, and in particular we cannot merely use bargaining solution $\mathbf{s}^{\alpha}$ prior to generation α because there are fewer than α patentholders. Third, we must recognize that the rate of quality improvements for the first α generations may differ from that for later generations.

In order to illustrate the basic effects of leading breadth, we consider the simple case where $\alpha = 2$ and $\mathbf{s}^2 = (0, 1)$. For this case, we only need to distinguish first-generation and second-generation products from later-generation products. Moreover, the reward to vertical R&D is the same for all generations, and so the rate of quality improvements is the same for all generations. In particular, the reward to vertical R&D is

$$V_v = \frac{\phi_v}{(\rho + \phi_v)^2}\left(\frac{LY}{B}\left(\frac{P^2 - 1}{P^2}\right)\right)$$

and the reward to horizontal R&D is

$$V_h = \frac{1}{\rho + \phi_v}\left(\frac{LY}{B}\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right)\right) + \frac{\phi_v}{(\rho + \phi_v)^2}\left(\frac{LY}{B}\left(\frac{\bar{\mu}P - 1}{\bar{\mu}P}\right)\right).$$

In the steady state, a proportion z_1 of industries will have first-generation products, a proportion z_2 of industries will have second-generation products, and a proportion $1 - z_1 - z_2$ of industries will have later-generation products, where $z_1 = g/(g + \phi_v)$ and $z_2 = g\phi_v/(\rho + \phi_v)^2$. As before, we make the simplifying assumptions that $z_1 \approx 0$ and $z_2 \approx 0$, in which case manufacturing labor per product is $n_m = LY/(BP^2)$, and economy-wide manufacturing labor is $N_m = Bn_m = LY/P^2$.[30] The resource constraint then requires $L = LY/P^2 + N_v + N_h$ at each date, or $Y = P^2(L - N_h - N_v)/L$. Recalling that $h^* = \lambda(P)V_v/V_h$ and that in the steady state $(L - N_h)/B = F(h^*)g/\left(\int_{h^*}^{\infty} h\, dF(h)\right)$, the steady-state outcome under any policy with $P \geq \gamma_o$ and $K = P^2$ will be (for the case when $n_v > 0$):

30 Once again, if we account for the endogeneity of z_1 and z_2, our qualitative conclusions are unchanged, but there may be additional feedback effects that alter magnitudes.

$$h^* = \lambda(P)\left(\frac{P^2 - 1}{P^2}\right) \Big/ \left[\left(\frac{\rho + \phi_v}{\phi_v}\right)\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right) + \left(\frac{\bar{\mu}P - 1}{\bar{\mu}P}\right)\right]$$

$$N_h/L = 1 - F(h^*)$$

$$n_v \text{ such that } \lambda(P)\frac{\phi_v}{(\rho + \phi_v)^2}\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\, dF(h)} - n_v\right)(P^2 - 1) = 1.$$

Now consider the effects of increasing leading breadth from $\alpha = 1$ to $\alpha = 2$. In our basic model, this policy change increases vertical R&D as long as ρ is small enough. To illustrate that a similar result holds here, consider the effects of this policy change in the limit case when $\rho \approx 0$. For this case, when $\alpha = 1$ the steady-state outcome involves

$$h^* = \lambda(P)\left(\frac{P - 1}{P}\right) \Big/ \left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right) \quad \text{and} \quad n_v = \left(\frac{P - 1}{P}\right)\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\, dF(h)}\right),$$

and when $\alpha = 2$ the steady-state outcome involves

$$h^* = \lambda(P)\left(\frac{P^2 - 1}{P^2}\right) \Big/ \left[\left(\frac{\bar{\mu} - 1}{\bar{\mu}}\right) + \left(\frac{\bar{\mu}P - 1}{\bar{\mu}P}\right)\right]$$

and

$$n_v = \left(\frac{P^2 - 1}{P^2}\right)\left(\frac{F(h^*)g}{\int_{h^*}^{\infty} h\, dF(h)}\right).$$

We can see that a policy change from $\alpha = 1$ to $\alpha = 2$ has two effects on vertical R&D n_v. First, increased leading breadth leads to increased market profits, and, if firms are patient enough, these increased profits lead to increased vertical R&D (as reflected by $(P^2 - 1)/P^2 > (P - 1)/P$). This effect is identical to that in our basic model. Second, increased leading breadth decreases the reward to vertical R&D relative to the reward to horizontal R&D, and therefore there is an increase in horizontal R&D (as reflected by a decrease in h^*), which leads to decreased vertical R&D. Because the second effect could be larger than the first, our result that, if people are patient enough, leading breadth stimulates vertical R&D does not necessarily survive in our non-scale growth model. In particular, our result survives if and only if the impact of increased leading breadth on horizontal R&D is sufficiently small.[31]

Acknowledgments

We thank Reto Foellmi, Gene Grossman, Vincenzo Denicolo, Suzanne Scotchmer, Chol-Won Li, Oded Galor, and an anonymous referee for useful comments. Tanja Zehnder

31 The fact that our patentability-requirement result unambiguously survives while our leading-breadth result might not survive is in part due to the fact that a patentability requirement can be increased by an incremental amount whereas leading breadth inherently involves discrete changes.

provided able research assistance. O'Donoghue would like to thank the France/Berkeley Fund, the Alfred P. Sloan Foundation, and the Institute for Advanced Studies, Vienna for financial support.

References

Aghion, P., and P. Howitt. (1992). ''A Model of Growth through Creative Destruction,'' *Econometrica* 60, 323–351.

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

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

Chang, H. (1995). ''Patent Scope, Antitrust Policy and Cumulative Innovation,'' *RAND Journal of Economics* 26, 34–57.

Cheng, L. K., and Z. Tao. (1999). ''The Impact of Public Policies on Innovation and Imitation: The Role of R&D Technology in Growth Models,'' *International Economic Review* 40, 187–207.

Chou, C., and O. Shy. (1993). ''The Crowding-Out Effects of Long Patent Durations,'' *RAND Journal of Economics* 24, 304–312.

Cohen, W. M., and R. C. Levin. (1989). ''Empirical Studies of Innovation and Market Structure.'' In R. Schmalensee, and R. D. Willig (eds.), *Handbook of Industrial Organization*, Amsterdam: Elsevier.

Cornelli, F., and M. Schankerman. (1999). ''Patent Renewals and R&D Incentives,'' *RAND Journal of Economics* 30, 197–213.

Davidson, C., and P. S. Segerstrom. (1993). ''Patent Enforcement and Economic Growth.'' FIEF Working Paper No. 113, Trade Union Institute for Economic Research, Stockholm.

Dinopoulos, E., and P. Thompson. (1998). ''Schumpeterian Growth without Scale Effects,'' *Journal of Economic Growth* 3, 313–335.

Gilbert, R. (2002). ''Antitrust for Patent Pools: A Century of Policy Evolution,'' Mimeo, U.C. Berkeley.

Gilbert, R., and C. Shapiro. (1990). ''Optimal Patent Length and Breadth,'' *RAND Journal of Economics* 21, 106–112.

Glass, A. (2001). ''Price Discrimination and Quality Improvement,'' *Canadian Journal of Economics* 34, 549–569.

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

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

Helpman, E. (1993). ''Innovation, Imitation, and Intellectual Property Rights,'' *Econometrica* 61, 1247–1280.

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

Hunt, R. (1999). ''Nonobviousness and the Incentive to Innovate: An Economic Analysis of Intellectual Property Reform.'' Working Paper 99-3, Federal Reserve Bank of Philadelphia.

Jones, C. I., (1995). ''Time Series Tests of Endogenous Growth Models,'' *Quarterly Journal of Economics* 110, 495–525.

Jones, C. I. (1999). ''Growth: With or Without Scale Effects,'' *American Economic Review, Papers and Proceedings* 89, 139–144.

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

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

Klemperer, P. (1990). ''How Broad Should the Scope of Patent Protection Be?'' *RAND Journal of Economics* 21, 113–130.

Klenow, P. J. (1996). ''Industry Innovation: Where and Why?'' *Carnegie-Rochester Conference Series on Public Policy* 44, 125–150.

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

Kremer, M. (1998). ''Patent Buyouts: A Mechanism for Encouraging Innovation,'' *Quarterly Journal of Economics* 113, 1137–1167.

Li, C.-W. (2000). ''Endogenous vs. Semiendogenous Growth in a Two-Sector-R&D Model,'' *Economic Journal* 110, C109–C122.

Li, C.-W. (2001). ''On the Policy Implications of Endogenous Technical Progress,'' *Economic Journal* 111, C164–C179.

Li, C.-W. (2002). ''Growth and Scale Effects: The Role of Knowledge Spillovers,'' *Economics Letters* 74, 177–185.

Matutes, C., P. Regibeau, and K. Rockett. (1996). ''Optimal Patent Design and the Diffusion of Innovations,'' *RAND Journal of Economics* 27, 60–83.

Merges, R. P., (1999). ''Institutions for Intellectual Property Transactions: The Case of Patent Pools,'' Mimeo, U.C. Berkeley.

Merges, R. P., and R. R. Nelson. (1990). ''On the Complex Economics of Patent Scope,'' *Columbia Law Review* 90, 839–916.

Nordhaus, W. (1969). *Invention, Growth, and Welfare*. Cambridge MA: MIT Press.

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

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

Peretto, P. (1998a). ''Technological Change, Market Rivalry, and the Evolution of the Capitalist Engine of Growth,'' *Journal of Economic Growth* 3, 53–80.

Peretto, P. (1998b). ''Technological Change and Population Growth,'' *Journal of Economic Growth* 3, 283–311.

Peretto, P., and S. Smulders. (1998). ''Specialization, Knowledge Dilution, and Scale Effects in an IO-based Growth Model,'' Mimeo, Duke University and Tilburg University.

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

Scotchmer, S. (1991). ''Standing on the Shoulders of Giants: Cumulative Research and the Patent Law,'' *Journal of Economic Perspectives* 5, 29–41.

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

Scotchmer, S. (1996b). ''Patents as an Incentive System.'' In Allen, B. (ed.), *Proceedings of the 1992 World Congress of the International Economic Association*, Cambridge MA: McMillan Press and St. Martin's Press.

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

Scotchmer, S., and J. Green. (1990). ''Novelty and Disclosure in Patent Law,'' *RAND Journal of Economics* 21, 131–146.

Segerstrom, P. (1992). ''On the Optimal Trade-Off Between Static and Dynamic Efficiency,'' Mimeo, Michigan State University.

Segerstrom, P. (1998). ''Endogenous Growth without Scale Effects,'' *American Economic Review* 88, 1290–1310.

Segerstrom, P., T. C. A. Anant, and E. Dinopoulos. (1990). ''A Schumpeterian Model of Product Life Cycle,'' *American Economic Review* 80, 1077–1091.

Smulders, S., and T. van der Klundert. (1995). ''Imperfect Competition, Concentration and Growth with Firm-Specific R&D,'' *European Economic Review* 39, 139–160.

Stokey, N. (1995). ''R&D and Economic Growth,'' *Review of Economic Studies* 62, 469–490.

Van Dijk, T. (1996). ''Patent Height and Competition in Product Improvements,'' *Journal of Industrial Economics* 44, 151–167.

Wright, B. (1983). ''The Economics of Invention Incentives: Patents, Prizes, and Research Contracts,'' *American Economic Review* 73, 691–707.

Young, A. (1998). ''Growth without Scale Effects,'' *Journal of Political Economy* 106, 41–63.

Zweimüller, J., and J. K. Brunner. (1998) ''Innovation and Growth with Rich and Poor Consumers,'' CEPR-Working Paper #1855.