Optimal Patent Breadth: Quantifying the Effects of Increasing Patent Breadth

Chu, Angus C.

University of Michigan

July 2007

Optimal Patent Breadth:

Quantifying the Effects of Increasing Patent Breadth

Angus C. Chu[*]

University of Michigan

July 2007

Abstract

In a generalized quality-ladder growth model, this paper firstly derives the optimal patent breadth and the socially optimal profit-sharing arrangement between patentholders. In this general-equilibrium setting, it identifies and derives a dynamic distortion of markup pricing on capital accumulation that has been neglected by previous studies on patent policy. Then, it quantitatively evaluates the effects of eliminating blocking patent and increasing patent breadth, and this exercise suggests a number of findings. Firstly, the market economy underinvests in R&D so long as a non-negligible fraction of long-run TFP growth is driven by R&D. Secondly, increasing patent breadth may be an effective solution to R&D underinvestment. The resulting effect on long-run consumption can be substantial because the harmful distortionary effects are relatively insignificant. However, the damaging effect of blocking patent arising from suboptimal profit-sharing arrangements between patentholders can be quantitatively significant. Finally, it considers the effect on consumption during the transition dynamics.

***Keywords*:** endogenous growth, intellectual property rights, patent breadth, R&D

***JEL classification*:** O31, O34

[*] Department of Economics, University of Michigan, 611 Tappan St., Ann Arbor, MI 48109, USA. Email address: anguscccc@umich.edu. I am grateful to Chris House, John Laitner, Dmitriy Stolyarov and Linda Tesar for invaluable advice and guidance and to Bob Barsky, Benjamin Chiao, Oli Coibion, Kathryn Dominquez, Yuriy Gorodnichenko, Osborne Jackson, David Levine, Doug Smith, Jing Zhang and seminar participants at the University of Michigan and at the Canadian Economics Association 2007 Conference for insightful comments and helpful suggestions. All errors are, of course, mine.

"Although length of protection is the most obvious policy lever for governing the profitability of an intellectual property right, it is not the instrument that shows up in patent disputes. Patent disputes almost always revolve around either validity of the patent in the first place, or the subtle question of how different another product must be in order not to infringe. The latter notion is colloquially known as the 'breadth' of the property right." – Scotchmer (2004, p. 103)

1. Introduction

Patent protection in the form of patent breadth has strengthened in the United States (US) since the 80's.[1] The facts that effective patent lives are very short and only a small fraction of patents are maintained until the end of the statutory term suggest that perhaps the term of patent is less of an important policy tool than patent breadth.[2] The first objective of this paper is to develop a tractable framework for a general-equilibrium analysis on optimal patent breadth using a quality-ladder growth model. In an environment with sequential innovations, patent breadth takes the form of lagging breadth and leading breadth. Lagging breadth provides patent protection against imitation while leading breadth provides patent protection against subsequent innovations, which might infringe existing patents. Whether an increase in *leading* breadth would enhance or dampen the incentives for research and development (R&D) depends on the profit-sharing arrangement between patentholders, who may engage in a complicated bargaining process. This paper derives the socially optimal profit-sharing arrangement between patentholders that should be implemented by the patent authority through designing an appropriate set of institutional policies for patent disputes to influence the patentholders' bargaining power. Upon enforcing this optimal profit-sharing arrangement, the optimal level of patent breadth can be determined by balancing the social marginal cost of distortions arising from patent protection and the social marginal benefit of R&D.

---

[1] See, e.g. Jaffe (2000), Gallini (2002), and Jaffe and Lerner (2004) for a comprehensive discussion. Hall, Jaffe and Trajtenberg (2002) provide data on the increase in the average number of patent citations, which can be viewed as a rough proxy for the broadening of patent breadth.

[2] See, e.g. O'Donoghue, Scotchmer and Thisse (1998) for a survey of empirical evidence on the short effective lifetime of patents and the small fraction of patents that are maintained until the end of the statutory term.

The second objective of this paper is to analyze the distortionary effects of increasing patent breadth. The patent-design literature emphasizes the tradeoff of patent protection between the incentives for R&D and the *static* distortionary effect of monopolistic markup pricing. However, mostly based on a partial-equilibrium setting, this literature neglects an important *dynamic* distortion on capital accumulation. In particular, increasing patent breath potentially raises the market value of patents on one hand and worsens the incentives for capital accumulation on the other by increasing the wedge between the marginal product of capital and its rental price. This paper identifies and analytically derives this distortionary effect in a generalized version of the quality-ladder growth model originating from Grossman and Helpman (1991) and Aghion and Howitt (1992).

The third objective of this paper is to provide a quantitative assessment on the effects of eliminating blocking patent and increasing patent breadth. Figure 1 shows that private spending on R&D in the US as a share of gross domestic product (GDP) has been rising sharply since the beginning of the 80's. Then, after a few years, the number of patents granted by the US Patent and Trademark Office also began to increase rapidly as shown in Figure 2. The calibration exercise takes as a premise the hypothesis that the incentive for private investments in R&D increases in response to broadening patent protection and makes use of the general-equilibrium framework to quantitatively evaluate the effects of increasing patent breadth and R&D along with the blocking-patent, static and dynamic distortionary effects on consumption. This numerical exercise suggests a number of findings. Firstly, the market economy underinvests in R&D relative to the first-best optimum so long as a non-negligible fraction of long-run total factor productivity (TFP) growth is driven by R&D. The quality-ladder growth model involves multiple externalities in R&D investment: (a) a negative intratemporal congestion or duplication externality; (b) a positive or negative externality in intertemporal knowledge spillover; (c) the monopolists' static surplus appropriability problem which is a positive externality; (d) the monopolists' dynamic surplus appropriability problem in the form of creative destruction which is also a positive externality; and (e) the business-stealing effect which is a negative externality. Furthermore, in the case of socially suboptimal profit-sharing arrangements between patentholders, there is an additional effect of blocking patent that reduces the incentives for R&D. Given the existence of positive and negative externalities, whether the market economy over- or under-invests in R&D depends mainly on the extents of intratemporal duplication and intertemporal spillover, which in turn are imputed from the balanced-growth condition between long-run TFP growth and R&D. Therefore, the larger is the fraction of long-run TFP growth driven by R&D, the more likely it is for the market economy to underinvest in R&D.

Secondly, increasing patent breadth may be an effective solution to the potential problem of R&D underinvestment. The resulting positive effect on long-run consumption can be substantial because the harmful effects of dynamic and static distortions are relatively insignificant. However, the damaging effect of blocking patent arising from socially suboptimal profit-sharing arrangements between patentholders can be quantitatively significant. Thirdly, the dynamic distortionary effect on capital accumulation has a more severe impact on consumption than the static distortionary effect from markup pricing unless the fraction of competitive industries in the economy is very large. Finally, it considers the effect on consumption during the transition dynamics. In particular, the economy does not always experience a significant fall in consumption in response to the increase in patent protection. Over a wide range of parameters, upon the strengthening of patent protection, consumption gradually rises towards the new balanced growth path by reducing physical investment and temporarily running down the capital stock. This finding contrasts that of Kwan and Lai (2003), whose model does not feature capital accumulation and hence predicts consumption losses during the transition path.

This paper relates to a number of studies. It provides an explanation and a potential solution to the R&D underinvestment problem identified by Jones and Williams (1998) and (2000). Jones and Williams (1998) develop a method to calculate the social rate of return to R&D based on endogenous-growth theory and show that estimates from the empirical productivity literature represent lower bounds on the true social rate of return. Using this information, they find that the socially optimal amount of R&D investment is at least two to four times larger than the actual amount. Jones and Williams (2000) adopt a different approach by calibrating a variety-expanding growth model to the data and obtain a similar conclusion that there is underinvestment in R&D over a wide range of parameters.[3] The current paper follows this latter approach by calibrating a generalized quality-ladder growth model with patent breadth as a policy instrument to show that the R&D underinvestment problem arises from insufficient patent breadth, and increasing patent breadth may be an effective solution to this problem. Furthermore, the calibration exercise takes into consideration Comin's (2004) critique that long-run TFP growth may not be solely driven by R&D.

In terms of qualitative analysis, it complements the patent-design literature,[4] which is mostly based on a partial-equilibrium setting, in providing a general-equilibrium analysis on optimal patent breadth and in identifying an important dynamic distortion on capital accumulation. O'Donoghue and Zweimuller (2004) is the first study that merges the patent-design and endogenous growth literatures to analyze the effects of patentability requirement, lagging and leading breadth on economic growth in a simple quality-ladder growth model. However, their focus was neither in characterizing the optimal patent breadth nor in quantifying the effects of eliminating blocking patent and increasing patent breadth. In addition, the current paper generalizes their model in a number of dimensions. For example, the usual Cobb-Douglas aggregator for intermediate goods is generalized to a CES aggregator to derive the condition under which patent breadth becomes ineffective in stimulating R&D. Goh and Olivier (2002) analyze the welfare effects of patent breadth in a two-sector variety-expanding growth model, and Grossman and Lai (2004) analyze the welfare effects of strengthening patent protection in developing countries as a result of the TRIPS agreement using a multi-country variety-expanding model. However, these studies do not analyze patent breadth in an environment with sequential innovations. Li (2001) analyzes the optimal policy mix of R&D subsidy and lagging breadth in a quality-ladder model with

[3] Stokey (1995) also calibrates an R&D-growth model to examine the range of parameters under which the market economy underinvests in R&D.

[4] The seminal work on optimal patent length is Nordhaus (1969). Some recent studies on optimal patent design include Tandon (1982), Gilbert and Shapiro (1990), Klemperer (1990), Green and Scotchmer (1995), O'Donoghue (1998), O'Donoghue, Scotchmer and Thisse (1998), Hunt (1999) and Scotchmer (2004). Judd (1985) provides the first dynamic general equilibrium analysis on optimal patent length.

endogenous step size, but he does not consider leading breadth. Furthermore, none of the abovementioned studies feature capital accumulation so that the dynamic distortion is absent.

Laitner (1982) identifies in an exogenous growth model with overlapping generations of households that the existence of an oligopolistic sector and its resulting pure profits as financial assets creates both the usual static distortion from markup pricing and an additional dynamic distortion on capital accumulation due to the crowding out of households' portfolio space, and he finds that the latter is more significant than the former. The current paper extends this study to show that this dynamic distortion also plays an important role and through a different channel in an R&D-driven endogenous growth model in which both patents and physical capital are owned by households as financial assets.

In terms of quantitative analysis, this paper relates to Kwan and Lai (2003) and Chu (2007). Kwan and Lai (2003) numerically evaluate the effects of extending the effective lifetime of patent in the variety-expanding model originating from Romer (1990) and find substantial welfare gains despite the temporary consumption losses during the transition path in their model. Chu (2007) uses a generalized variety-expanding model and finds that whether or not an extension in the patent length is effective in stimulating R&D depends crucially on the patent-value depreciation rate. At the empirical range of patent-value depreciation rates estimated by previous studies, patent extension has only limited effects on R&D and thus social welfare. Therefore, Chu (2007) and the current paper together provide a comparison on the effectiveness of patent length and patent breadth in solving the R&D underinvestment problem. The crucial difference between these two policy instruments arises because patent length affects future monopolistic profits while patent breadth affects current monopolistic profits.

The rest of the paper is organized as follows. Section 2 describes the generalized quality-ladder model and derives the analytical characterization of optimal patent breadth and the dynamic distortion on capital accumulation. Section 3 calibrates the model and numerically evaluates the effects of eliminating blocking patent and increasing patent breadth on consumption. The final section concludes with some important caveats. Appendix I contains the proofs.

2. Optimal Patent Breadth

The patent-design literature has identified and analyzed four patent-policy tools: (a) the term of patent or simply patent length; (b) patentability requirement; (c) lagging breadth; and (d) leading breadth.[5] In a standard quality-ladder growth model, lagging breadth (i.e. patent protection against imitation) is assumed to be complete while leading breadth (i.e. patent protection against subsequent innovations) is assumed to be zero. This section derives the second-best optimal level of lagging and leading breadth chosen by a benevolent government in a generalized quality-ladder growth model.

The model is a generalized version of Grossman and Helpman (1991) and Aghion and Howitt (1992). To prevent the model from overestimating the social benefits of R&D and hence the extents of underinvestment in R&D, long-run TFP growth is assumed to be driven by both R&D investment and an exogenous process as in Comin (2004). To prevent the model from overstating the effectiveness of patent breadth in stimulating R&D, the usual Cobb-Douglas aggregator for the quality-enhancing intermediate goods is generalized to a CES aggregator as in Laitner and Stolyarov (2005). To maintain the analytical tractability of the aggregate conditions under the CES aggregator, all the intermediate-goods industries are assumed to be monopolistic in this section; consequently, the static distortion is absent. To introduce the static distortionary effect of markup pricing into the model, a special case of the Cobb-Douglas aggregator will be considered in addition to the CES aggregator when performing the numerical exercises in Section 3. Furthermore, computation of the transition dynamics is possible under the Cobb-Douglas aggregator.[6] In order to perform a more realistic calibration, the model is further modified to include physical capital, which is a factor input for the production of intermediate goods and R&D, and the final goods can be used for consumption or investment in capital. Finally, the class of first-generation R&D-driven endogenous growth models, such as Grossman and Helpman (1991) and Aghion and Howitt

---

[5] See, e.g. O'Donoghue and Zweimuller (2004) for an overview of these four patent-policy tools. For a more detailed discussion on patentability requirement and leading breadth, refer to O'Donoghue (1998) and O'Donoghue, Scotchmer and Thisse (1998).

[6] Although the arrival rate of innovations varies along the transition path, a tractable form for the law of motion for aggregate technology can still be derived under the Cobb-Douglas aggregator but not under the CES aggregator.

(1992), exhibits scale effects and is inconsistent with the empirical evidence in Jones (1995a).[7] In the present model, scale effects are eliminated by assuming increasing difficulty in R&D successes as in Segerstrom (1998), which becomes a semi-endogenous growth model.[8]

The various components of the model are presented in Sections 2.1–2.9, and the balanced-growth equilibrium is defined in Section 2.10. Section 2.11 derives the first-best social optimum, and Section 2.12 characterizes the second-best optimal level of patent breadth.

2.1. Representative Household

The infinitely-lived representative household maximizes life-time utility that is a function of per-capita consumption c_t of the numeraire final goods and is assumed to have the iso-elastic form given by

(1) $$U = \int_0^\infty e^{-(\rho-n)t} \frac{c_t^{1-\sigma}}{1-\sigma} dt .$$

$\sigma \geq 1$ is the inverse of the elasticity of intertemporal substitution. The household has $L_t = L_0 \exp(nt)$ members at time *t*. The population size at time 0 is normalized to one, and $n > 0$ is the exogenous population growth rate. ρ is the subjective discount rate. To ensures that lifetime utility is bounded,

(A1) $$\rho > n .$$

The household maximizes (1) subject to a sequence of budget constraints given by

(2) $$\dot{a}_t = a_t(r_t - n) + w_t - c_t .$$

Each member of the household inelastically supplies one unit of homogenous labor in each period to earn a real wage income w_t. a_t is the value of risk-free financial assets in the form of patents and physical

[7] See, e.g. Jones (1999) for an excellent theoretical analysis on scale effects.

[8] In a semi-endogenous growth model, the balanced-growth rate is determined by the exogenous population growth rate. An increase in the share of R&D factor inputs raises the *level* of the balanced growth path while holding the balanced growth rate constant. Since increasing R&D has no long-run growth effect in this model, the estimated effects on consumption are likely to be more conservative than in other fully endogenous growth models.

capital owned by each household member, and r_t is the real rate of return on these assets. The familiar Euler equation derived from the intertemporal optimization is

$$\dot{c}_t = c_t(r_t - \rho)/\sigma . \tag{3}$$

Along the balanced-growth path, c_t increases at a constant rate g_c. The steady-state real interest rate is

$$r = \rho + g_c\sigma . \tag{4}$$

2.2. Final Goods

This sector is characterized by perfect competition, and the producers take both the output price and input prices as given. The production function for the final goods Y_t is a CES aggregator of a continuum of differentiated quality-enhancing intermediate goods $X_t(j)$ for $j \in [0,1]$ given by

$$Y_t = \left(\int_0^1 X_t^{\varepsilon}(j)dj \right)^{1/\varepsilon} , \tag{5}$$

where $\varepsilon \in [0, 1)$. The constant elasticity of substitution as well as the absolute value of demand elasticity is $1/(1-\varepsilon)$. This formulation nests the usual Cobb-Douglas aggregator in quality-ladder models as a special case with $\varepsilon = 0$. The familiar aggregate price index is $P_t = \left(\int_0^1 P_t^{\varepsilon/(\varepsilon-1)}(j)dj \right)^{(\varepsilon-1)/\varepsilon} = 1$, and the demand curve for each variety of intermediate goods is

$$X_t(j) = P_t(j)^{-1/(1-\varepsilon)} Y_t . \tag{6}$$

2.3. Intermediate Goods

There is a continuum of monopolistic industries producing the differentiated quality-enhancing intermediate goods $X_t(j)$ for $j \in [0,1]$, and each industry is dominated by a temporary industry leader, who owns the latest R&D-driven technology for production. The production function in each industry *j* has constant returns to scale in labor and capital inputs and is given by

(7) $$X_t(j) = z^{m_t(j)} Z_t K_{x,t}^{\alpha}(j) L_{x,t}^{1-\alpha}(j).$$

$K_{x,t}(j)$ and $L_{x,t}(j)$ are respectively the capital and labor inputs for producing intermediate-goods *j* at time *t*. $Z_t = Z_0 \exp(g_Z t)$ represents an exogenous process of productivity improvement that is common across all industries and is freely available to all producers. $z^{m_t(j)}$ is the industry leader's level of R&D-driven technology, which is increasing over time through R&D investment and successful innovations. $z > 1$ is the exogenous step-size of a technological improvement arising from each innovation. $m_t(j)$, which is an integer, is the number of innovations that has occurred in industry *j* as of time *t*. The marginal cost of production in industry *j* is

(8) $$MC_t(j) = \frac{1}{z^{m_t(j)} Z_t} \left(\frac{R_t}{\alpha} \right)^{\alpha} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha},$$

where R_t is the rental price of capital. The optimal price is a constant markup μ over the marginal cost of production given by

(9) $$P_t(j) = \mu MC_t(j).$$

The profit-maximizing markup for an unconstrained monopolist is $1/\varepsilon$. With a Cobb-Douglas aggregator, $\varepsilon = 0$; therefore, it is always the closest rival's marginal cost that is the binding constraint. Then, in the standard case of *complete lagging breadth* and *zero leading breadth* (to be defined in Sections 2.5 and 2.6), the industry leader is able to charge a markup of z over the marginal cost without losing its market share to the closest rival. With a CES aggregator, $\mu \equiv \min\{z, 1/\varepsilon\}$. If $z \geq 1/\varepsilon$, then increasing leading breadth would have no stimulating effects on R&D because the industry leader would always choose a markup of $1/\varepsilon$ regardless the level of leading breadth. To analyze the implications of increasing leading breadth, the following parameter condition is assumed for the theoretical analysis

(A2) $$z < 1/\varepsilon.$$ [9]

Then, the amount of profit earned by the leader of industry j at time t is

(10) $$\pi_t(j) = (z-1)MC_t(j)X_t(j).$$

2.4. Patent Breadth

This subsection presents the Bertrand equilibrium price and profit in the presence of patent breadth, which is denoted by η, under the optimal profit-sharing arrangement. Then, in the following two subsections, η is decomposed into lagging breadth η_{lag} and leading breadth η_{lead} (i.e. $\eta = \eta_{lag} + \eta_{lead}$) to demonstrate the underlying assumptions behind the following analytically tractable expressions

(11) $$P_t(j) = z^{\eta} MC_t(j)$$

(12) $$\pi_t(j) = (z^{\eta} - 1)MC_t(j)X_t(j)$$

for $\eta \in (0,\infty)$ and $j \in [0,1]$. The expression for the equilibrium price is consistent with the seminal work of Gilbert and Shapiro's (1990) interpretation of "breadth as the ability of the patentee to raise price." A broader patent breadth corresponds to a larger η, and vice versa. Therefore, an increase in patent breadth enhances the incentives for R&D by raising the amount of monopolistic profit captured by each innovation but worsens the distortionary effects of markup pricing. This discussion implicitly assumes that $z^{\eta} < 1/\varepsilon$ because the markup is now given by $\mu \equiv \min\{z^{\eta}, 1/\varepsilon\}$. When $z^{\eta} = 1/\varepsilon$, patent breadth has no more stimulating effects on R&D and no more distortionary effects from markup-pricing.

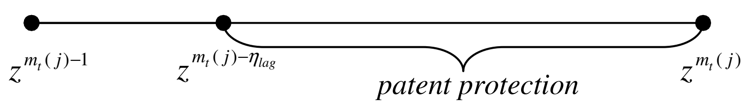
2.5. Lagging Breath

The first deviation from standard quality-ladder models is the introduction of incomplete lagging breadth. Assume zero leading breadth $\eta_{lead} = 0$ as in standard models for now. To reiterate, each innovation is a

[9] According to Broda and Weinstein (2006), the elasticity of substitution $1/(1-\varepsilon)$ for differentiated goods has a mean estimate of 4.7-5.2 and a median estimate of 2.1-2.5. Given an empirical markup of 1.10 (e.g. Laitner and Stolyarov (2004)), it seems reasonable to assume that the markup in the data is not determined by demand elasticity.

constant step-size z of a technological improvement, and this production technology, once invented, becomes public knowledge to fulfill the disclosure requirement for obtaining a patent. In the case of complete lagging breadth, the patent for $m_t(j)$ allows the new industry leader to produce with any technology level $\in (z^{m_t(j)-1}, z^{m_t(j)}]$, but the profit-maximizing level is $z^{m_t(j)}$. The former industry leader, who holds the patent for $m_t(j)-1$, is now also technologically feasible to upgrade its production process. However, to do so, she would infringe the patent of the new industry leader, and any licensing agreement would drive the licensee's profit to zero.

The parameter $\eta_{lag} \leq 1$ represents the degree of lagging breadth. In the special case of complete lagging breadth $\eta_{lag} = 1$, any technology level beyond $z^{m_t(j)-1}$ is protected by the patent for $m_t(j)$. In the case of incomplete lagging breadth $\eta_{lag} < 1$, only technology level beyond $z^{m_t(j)-\eta_{lag}}$ is protected. The following diagram illustrates the concept of incomplete lagging breadth.



In other words, although the invention is a quality improvement of z, the patent only protects part of this invention $z^{\eta_{lag}}$ against imitators. Therefore, with incomplete lagging breadth, the Bertrand equilibrium price becomes

(13) $$P_t(j) = z^{\eta_{lag}} MC_t(j)$$

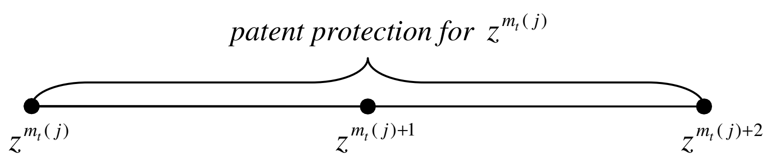
for $\eta_{lag} \in (0,1)$ and $j \in [0,1]$. The amount of profit is

(14) $$\pi_t(j) = (z^{\eta_{lag}} - 1) MC_t(j) X_t(j)$$

for $\eta_{lag} \in (0,1)$ and $j \in [0,1]$. Incomplete protection against imitators forces the industry leader to lower its markup. On one hand, incomplete lagging breadth reduces the distortionary effects of markup pricing; on the other hand, the reduced profit worsens the incentives for R&D.

2.6. Leading Breadth

The second deviation from standard models is the introduction of nonzero leading breadth, which protects patentholders against subsequent innovators. The formulation of leading breadth originates from O'Donoghue and Zweimuller (2004). Assume complete lagging breadth $\eta_{lag} = 1$ as in standard models for now. The degree of leading breadth is represented by $\eta_{lead} \in \{0,1,2,...\}$. Standard models assume zero leading breadth (i.e. $\eta_{lead} = 0$). For example, if $\eta_{lead} = 1$, then the most recent innovation infringes the patent of the second-most recent innovator. If $\eta_{lead} = 2$, then the most recent innovation infringes the patents of the second-most and the third-most recent innovators, etc. The following diagram illustrates the concept of nonzero leading breadth with an example of leading breadth equal two.



Therefore, nonzero leading breadth facilitates the new industry leader and the previous innovators, whose patents are infringed, to consolidate market power through licensing agreements resulting in a higher markup.[10] The Bertrand equilibrium price under nonzero leading breadth is

(15) $$P_t(j) = z^{1+\eta_{lead}} MC_t(j)$$

for $\eta_{lead} \in \{1,2,...\}$ and $j \in [0,1]$. Assumption 1 is *sufficient* to derive this equilibrium markup price.

Assumption 1: *An infringed patentholder cannot become the next industry leader while she is still covered by a licensing agreement in that industry.*[11]

[10] See, e.g. Gallini (2002) and O'Donoghue and Zweimuller (2004), for a discussion on market-power consolidation through licensing agreements.

[11] The sufficiency of this assumption in determining the markup price is most easily understood with an example. Suppose leading breadth is one and lagging breadth is complete, the lower bound on the profit-maximizing markup is the square of z, which is the limit price from the collusion of the most recent and the second-most recent innovators against the third-most recent innovator, whose patent is not infringed upon by the most recent invention. In this example, the limit-pricing markup would be even larger if the third-most recent innovator happens to be the new industry leader. Continuing this reasoning, the markup could grow without bound or until it equals $1/\varepsilon$;

Then, the amount of monopolistic profit generated in each period by the licensing agreement between the industry leader and the infringed patentholders is

(16) $$\pi_t(j) = (z^{1+\eta_{lead}} - 1)MC_t(j)X_t(j)$$

for $\eta_{lead} \in \{1,2,...\}$ and $j \in [0,1]$.

Determining the share of profit obtained by the industry leader requires solving the profit-sharing arrangement (i.e. the terms in the licensing agreement) between patentholders. A stationary outcome is assumed to simplify the analysis.

Assumption 2: *There is a set of stationary profit-sharing arrangements for each* $\eta_{lead} \in \{1,2,...\}$ *denoted by* $\sigma^{\eta_{lead}} = (\sigma_1^{\eta_{lead}}, \sigma_2^{\eta_{lead}}, ..., \sigma_{1+\eta_{lead}}^{\eta_{lead}}) \in [0,1]$, *where* $\sigma_i^{\eta_{lead}}$ *is the share of profit received by the i-th most recent innovator when leading breadth is* η_{lead}, *and* $\sum_{i=1}^{1+\eta_{lead}} \sigma_i^{\eta_{lead}} = 1$.

Although the shares of profits and licensing fees eventually received by the owner of an invention are constant overtime, the present value of profits is determined by the actual profit-sharing arrangement. The two extreme cases are: (a) *complete frontloading* $\sigma^{\eta_{lead}} = (1,0,...,0)$; and (b) *complete backloading* $\sigma^{\eta_{lead}} = (0,0,...,1)$. Complete frontloading maximizes the stimulating effect of leading breadth on R&D by maximizing the present value of profits. The opposite effect of blocking patent arises when profits are backloaded, and complete backloading maximizes this damaging effect on the incentives for R&D.

Assumptions 1 and 2 originate from O'Donoghue and Zweimuller (2004) to formalize the modeling of leading breadth, but they did not derive the equilibrium outcome that requires solving the bargaining game between patentholders. Instead, this paper characterizes the optimal patent breadth by deriving the optimal profit-sharing arrangement.

---

therefore, Assumption 1 is made to rule out this possibility. The empirical plausibility of this assumption is appealed to the existence of antitrust policy.

Assumption 3: *The patent authority is able to enforce the socially optimal profit-sharing arrangement in the licensing agreements between patentholders.*

Proposition 1a: *For any given level of patent breadth, the complete frontloading profit-sharing arrangement is socially optimal if there is underinvestment in R&D in the decentralized equilibrium.*

Proposition 1b: *In the special case of labor being the only factor input for R&D, the complete frontloading profit-sharing arrangement is socially optimal if and only if there is underinvestment in R&D in the decentralized equilibrium.*

Intuitively, the first-order distortionary effect of markup pricing is determined by η, independent of the profit-sharing arrangement between patentholders. Therefore, given a level of patent breadth, the society is better off by having a profit-sharing arrangement that creates the largest incentives for R&D if there is underinvestment in R&D in the market economy. However, the underinvestment in R&D is not a necessary condition when capital is also a factor input for R&D because stimulating the incentives for R&D in this case also increases the rate of investment in capital that partly offsets the dynamic distortionary effect of markup pricing on capital accumulation.

Proposition 1 establishes the condition for the social optimality of the complete frontloading profit-sharing arrangement, in which the infringed patentholders of previous inventions allow the new industry leader to capture the entire amount of profits from her invention until the next innovation occurs. Every successful innovator goes through the cycle of being an infringing industry leader initially and an infringed patentholder subsequently. Therefore, the distinction between the frontloading and backloading of profits matters only because the real interest rate is higher than the profit growth rate. A real-world example of this profit-sharing arrangement is a royalty-free cross-licensing agreement.[12] From a policy perspective, the complete frontloading profit-sharing arrangement should be implemented by the patent authority through the following policies: (a) compulsory licensing with an upper limit on the amount of licensing fees charged to subsequent inventors of more advanced technology; and (b) making patent-infringement cases in court favorable to subsequent inventors of more advanced technology.

Given Proposition 1, the equilibrium price and the amount of profit for an industry leader are respectively $P_t(j) = z^{\eta} MC_t(j)$ and $\pi_t(j) = (z^{\eta} - 1) MC_t(j) X_t(j)$. In the case of *complete* lagging breadth and *zero* leading breadth, $\eta = 1$. In the case of *incomplete* lagging breadth and *zero* leading breadth, $\eta \in (0,1)$. In the case of *complete* lagging breadth and *nonzero* leading breadth, $\eta \in \{2,3,...\}$. In the general case of *incomplete* lagging breadth and *nonzero* leading breadth, $\eta \in (0,\infty)$. For example, $\eta = 1.5$ corresponds to lagging breadth of 0.5 and leading breadth of 1.

2.7. Aggregation

The aggregate production function for the final goods is

(17) $$Y_t = \left(\int_0^1 X_t^{\varepsilon}(j) dj \right)^{1/\varepsilon} = A_t Z_t K_{x,t}^{\alpha} L_{x,t}^{1-\alpha},$$

where $A_t \equiv \left(\int_0^1 (z^{m_t(j)})^{\varepsilon/(1-\varepsilon)} dj \right)^{(1-\varepsilon)/\varepsilon}$ is the level of R&D-driven technology. $K_{x,t} = \int_0^1 K_{x,t}(j) dj$ and $L_{x,t} = \int_0^1 L_{x,t}(j) dj$ are total labor and capital inputs for production. The market-clearing condition for the final goods is

(18) $$Y_t = C_t + I_t.$$

[12] Under a cross-licensing agreement, each company lists a large number of patents that it owns and the companies are allowed to use any of the patents listed in the agreement. If the companies' portfolios are similar in size and quality, the agreement may involve no monetary compensation. See, e.g. Jaffe and Lerner (chapter 2, 2004).

$C_t = L_t c_t$ is the aggregate consumption, and I_t is the investment in physical capital. The factor payments for the final goods are

(19) $$Y_t = w_t L_{x,t} + R_t K_{x,t} + \pi_t .$$

$\pi_t = \int_0^1 \pi_t(j)dj$ is the total amount of monopolistic profits. Substituting (7) and (8) into (12) and then summing over all industries yields

(20) $$\pi_t = \left(\frac{\mu - 1}{\mu}\right) Y_t .$$

Therefore, the growth rate of monopolistic profits equals the growth rate of output denoted by g_Y. The amount of factor payments for labor and capital inputs are

(21) $$w_t L_{x,t} = \left(\frac{1-\alpha}{\mu}\right) Y_t ,$$

(22) $$R_t K_{x,t} = \left(\frac{\alpha}{\mu}\right) Y_t .$$

(22) shows that the markup drives a wedge between the marginal product of capital and its rental price. As will be shown below, this wedge creates a distortion on the rate of investment in physical capital.

Finally, the value of GDP *should* include the amount of investment in R&D such that

(23) $$GDP_t = Y_t + w_t L_{r,t} + R_t K_{r,t} .$$ [13]

$L_{r,t}$ and $K_{r,t}$ are respectively the number of workers and the amount of capital for R&D.

2.8. Capital Accumulation

The market-clearing condition for physical capital is

[13] In the national income account, R&D investment is treated as an expenditure on intermediate goods. Therefore, the values of investment and GDP in the data are I_t and Y_t respectively. The Bureau of Economic Analysis and the National Science Foundation's R&D satellite account provides preliminary estimates on the effects of including R&D as an intangible asset in the national income accounts.

(24) $$K_t = K_{x,t} + K_{r,t}\,.$$

K_t is the total amount of capital available in the economy at time t. The law of motion for capital is

(25) $$\dot{K}_t = I_t - K_t\delta$$

δ is the rate of depreciation. Denote the balanced-growth rate of capital by g_K; then, the endogenous steady-state investment rate in physical capital is

(26) $$i = (g_K + \delta)K_t / Y_t$$

for all *t*. The no-arbitrage condition $r_t = R_t - \delta$ for the holding of capital and (22) imply that the steady-state capital-output ratio is

(27) $$\frac{K_t}{Y_t} = \frac{\alpha}{\mu(1-s_K)(r+\delta)}\,.$$

s_K is the endogenous steady-state share of capital for R&D. Substituting (27) into (26) yields

(28) $$i = \frac{\alpha}{\mu(1-s_K)}\left(\frac{g_K+\delta}{r+\delta}\right).$$

In the Romer model, (skilled) labor is the only factor input for R&D (i.e. $s_K = 0$); therefore, the distortionary effect of markup pricing on the rate of investment is unambiguously negative (i.e. $\partial i / \partial \mu < 0$). In the current model, there is an opposing positive effect operating through s_K. Intuitively, an increase in patent breadth raises the private return on R&D and consequently, the share of capital employed in the R&D sector. Proposition 2 in Section 2.11 shows that the negative distortionary effect still dominates if the intermediate-goods sector is at least as capital intensive as the R&D sector.

2.9. R&D

$V_t(j)$ is the value of the patent owned by the leader in industry *j* at time *t* and is determined by the following no-arbitrage condition

(29) $$r_t V_t(j) = \pi_t(j) + \dot{V}_t(j) - \lambda_t V_t(j)\,.$$

The first terms in the right is the flow profit generated by the patent at time *t*. The second term is the capital gain due to the growth in profit. The third term is the expected value of capital loss due to creative destruction, and λ_t is the Poisson arrival rate of the next innovation in the same industry. This no-arbitrage condition can be re-expressed as

(30)
$$V_t(j) = \frac{\pi_t(j)}{r_t + \lambda_t - \dot{V}_t(j)/V_t(j)}.$$

The aggregate value of the patents owned by all the industry leaders at time *t* is

(31)
$$V_t = \int_0^1 V_t(j)dj = \left(\frac{\mu - 1}{\mu}\right)\frac{Y_t}{r_t + \lambda_t - \dot{V}_t/V_t}.$$

Since the amount of monopolistic profits varies across industries with the CES aggregator, it leads to strategic considerations in terms of targeting innovations to a particular industry. To avoid this problem, the following assumption is made.

Assumption 4: *Innovation successes of the R&D entrepreneurs are randomly assigned to the industries in the intermediate-goods sector.*

Therefore, the steady-state no-arbitrage value of achieving a new successful innovation at time *t* is the expected present value of the stream of monopolistic profits given by

(32)
$$V_t = \left(\frac{\mu - 1}{\mu}\right)\frac{Y_t}{r + \lambda - g_Y}.$$[14]

The arrival rate of an innovation success for an R&D entrepreneur $h \in [0,1]$ is a function of labor input $L_{r,t}(h)$ and capital input $K_{r,t}(h)$ given by

(33)
$$\lambda_t(h) = \overline{\varphi}_t K_{r,t}^{\beta}(h) L_{r,t}^{1-\beta}(h).$$[15]

[14] Because λ is pinned down by the population growth rate along the balanced growth path (to be shown below), the value of a patent is unambiguously increasing in η. This implication is consistent with the empirical finding in Lerner (1994) that patent breadth is positively correlated with the market value in his sample of biotechnology firms.

$\overline{\varphi}_t$ is a productivity parameter that the entrepreneurs take as given. The amount of expected profit from R&D is

(34) $$E_t[\pi_{r,t}(h)] = V_t\lambda_t(h) - w_tL_{r,t}(h) - R_tK_{r,t}(h).$$

The first-order conditions are

(35) $$(1-\beta)V_t\overline{\varphi}_t(K_{r,t}(h)/L_{r,t}(h))^{\beta} = w_t,$$

(36) $$\beta V_t\overline{\varphi}_t(K_{r,t}(h)/L_{r,t}(h))^{\beta-1} = R_t.$$

To eliminate scale effects and capture various externalities, the *individual* R&D productivity parameter $\overline{\varphi}_t$ at time t is assumed to be decreasing in the level of R&D-driven technology A_t such that

(37) $$\overline{\varphi}_t = \frac{\varphi(K_{r,t}^{\beta}L_{r,t}^{1-\beta})^{\gamma-1}}{A_t^{1-\phi}},$$

where $K_{r,t} = \int_0^1 K_{r,t}(h)dh$ and $L_{r,t} = \int_0^1 L_{r,t}(h)dh$. $\gamma \in (0,1]$ captures the intratemporal negative congestion or duplication externality or the so-called "stepping on toes" effects, and $\phi \in (-\infty,1)$ captures the externality of intertemporal knowledge spillovers.[16] Given that the arrival of innovations follows a Poisson process, Laitner and Stolyarov (2005) appeal to the Law of Large Numbers to show that the aggregate technology can be re-expressed as $A_t = \exp\left(\lambda\tilde{z}(\varepsilon)\,t\right)$, where $\tilde{z}(\varepsilon) \equiv (z^{\varepsilon/(1-\varepsilon)} - 1)(1-\varepsilon)/\varepsilon$. Therefore, the law of motion for R&D-driven technology along the balanced-growth path, in which λ is constant, is given by

(38) $$\dot{A}_t = A_t\lambda\tilde{z}(\varepsilon) = A_t\overline{\varphi}_tK_{r,t}^{\beta}L_{r,t}^{1-\beta}\tilde{z}(\varepsilon) = A_t^{\phi}(K_{r,t}^{\beta}L_{r,t}^{1-\beta})^{\gamma}\varphi\tilde{z}(\varepsilon).$$

---

[15] This specification nests the "knowledge-driven" specification in Romer (1990) as a special case with $\beta = 0$ and the "lab equipment" specification in River-Batiz and Romer (1991) as a special case with $\beta = \alpha$.

[16] This specification captures how semi-endogenous growth models eliminate scale effects as in Jones (1995b). $\phi \in (0,1)$ corresponds to the "standing on shoulder" effect, in which the *economy-wide* R&D productivity $A_q\overline{\varphi}$ increases as the level of R&D-driven technology increases (see the law of motion for R&D-driven technology). On the other hand, $\phi \in (-\infty,0)$ corresponds to the "fishing out" effect, in which early technology is relatively easy to develop and $A_q\overline{\varphi}$ decreases as the level of R&D-driven technology increases.

Along the balanced-growth path, the growth rate of R&D-driven technology denoted by g_A is related to the population growth rate such that

(39)
$$g_A = \frac{\dot{A}_t}{A_t} = \frac{(K_{r,t}^{\beta} L_{r,t}^{1-\beta})^{\gamma}}{A_t^{1-\phi}} \varphi \tilde{z}(\varepsilon) = \left(\frac{\gamma\beta}{1-\phi}\right) g_K + \left(\frac{\gamma(1-\beta)}{1-\phi}\right) n .$$

Then, the steady-state rate of creative destruction is $\lambda = g_A / \tilde{z}(\varepsilon)$.

2.10. Balanced-Growth Equilibrium

The analysis starts at $t = 0$ when the economy has reached its balanced-growth path corresponding to the patent policy $\{\eta\}$. The equilibrium is a sequence of prices $\{w_t, r_t, R_t, P_t(j), V_t\}_{t=0}^{\infty}$ and a sequence of allocations $\{a_t, c_t, I_t, Y_t, X_t(j), K_{x,t}(j), L_{x,t}(j), K_{r,t}(h), L_{r,t}(h), K_t, L_t\}_{t=0}^{\infty}$ that are consistent with the initial conditions $\{K_0, L_0, Z_0, A_0, \overline{\varphi}_0\}$ and their subsequent laws of motions. Also, in each period,

(a) the representative household chooses $\{a_t, c_t\}$ to maximize utility taking $\{w_t, r_t\}$ as given;

(b) the competitive firms in final-goods sector choose $\{X_t(j)\}$ to maximize profits according to the production function taking $\{P_t(j)\}$ as given;

(c) the industry leaders $j \in [0,1]$ in the intermediate-goods sector choose $\{P_t(j), K_{x,t}(j), L_{x,t}(j)\}$ to maximize profits according to the Bertrand price competition and the production function taking $\{R_t, w_t\}$ as given;

(d) the entrepreneurs $h \in [0,1]$ in the R&D sector choose $\{K_{r,t}(h), L_{r,t}(h)\}$ to maximize profits according to the R&D production function taking $\{\overline{\varphi}_t, V_t, R_t, w_t\}$ as given;

(e) the market for the final-goods clears such that $Y_t = C_t + I_t$;

(f) the full employment of capital such that $K_t = K_{x,t} + K_{r,t}$; and

(g) the full employment of labors such that $L_t = L_{x,t} + L_{r,t}$.

Equating the first-order conditions (21) and (35) and imposing the balanced-growth condition

(40) $$g_A = \overline{\varphi}_t L_{r,t}^{1-\beta} K_{r,t}^{\beta} \tilde{z}(\varepsilon)$$

yield the steady-state R&D share of labor inputs given by

(41) $$\frac{s_L}{1-s_L} = \frac{1-\beta}{1-\alpha}\left(\frac{(\mu-1)\lambda}{r+\lambda-g_Y}\right).$$

Similarly, solving (22), (36) and (40) yields the steady-state R&D share of capital inputs given by

(42) $$\frac{s_K}{1-s_K} = \frac{\beta}{\alpha}\left(\frac{(\mu-1)\lambda}{r+\lambda-g_Y}\right).$$

The balanced-growth rates of various variables are given as follows. Given that the steady-state investment rate is constant, the steady-state growth rate of per capita consumption is

(43) $$g_c = g_Y - n\,.$$

From the aggregate production function (17), the steady-state growth rates of output and capital are

(44) $$g_Y = g_K = n + (g_A + g_Z)/(1-\alpha)\,.$$

Using (39) and (44), the steady-state growth rate of R&D-driven technology is determined by the exogenous population growth rate n and productivity growth rate g_Z given by

(45) $$g_A = \left(\frac{1-\phi}{\gamma} - \frac{\beta}{1-\alpha}\right)^{-1}\left(n + \frac{\beta}{1-\alpha} g_Z\right).$$

Long-run TFP growth denoted by $g_{TFP} \equiv g_A + g_Z$ is empirically observed. For a given g_{TFP}, a higher value of g_Z implies a lower value of g_A as well as a lower calibrated value for $\gamma/(1-\phi)$ indicating smaller social benefits from R&D.

2.11. First-Best Social Optimum

To derive the socially optimal equilibrium rate of investment i^* and R&D shares of labor s_L^* and capital s_K^*, the social planner maximizes

(46) $$U = \int_0^\infty e^{-(\rho-n)t}\, \frac{((1-i)Y_t / L_t)^{1-\sigma}}{1-\sigma}\, dt$$

subject to: (a) the aggregate production function expressed in terms of s_L and s_K given by

(47) $$Y_t = A_t Z_t (1-s_K)^{\alpha} (1-s_L)^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha};$$

(b) the law of motion for capital expressed in terms of i given by

(48) $$\dot{K}_t = iY_t - K_t\delta;$$

and (c) the law of motion for R&D-driven technology expressed in terms of s_L and s_K given by

(49) $$\dot{A}_t = A_t^{\phi} (s_K)^{\beta\gamma} (s_L)^{(1-\beta)\gamma} K_t^{\beta\gamma} L_t^{(1-\beta)\gamma} \varphi \tilde{z}(\varepsilon).$$

After solving this maximization problem, the modified Golden-rule rate of investment is

(50) $$i^* = \left(\alpha + \beta \frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} \right) \frac{g_K + \delta}{\rho + g_c\sigma + \delta}.$$

Proposition 2 provides the condition under which the markup-pricing distortion moves the market equilibrium rate of investment i away from the social optimum i^*.

Proposition 2a: *The decentralized equilibrium rate of investment is below the socially optimal investment rate if either there is underinvestment in R&D or labor is the only factor input for R&D.*

Proposition 2b: *An increase in patent breadth leads to a reduction in the decentralized equilibrium rate of investment if the intermediate-goods sector is at least as capital intensive as the R&D sector.*

Similarly, the socially optimal R&D shares of labor s_L^* and capital s_K^* are respectively

(51) $$\frac{s_L^*}{1-s_L^*} = \frac{1-\beta}{1-\alpha} \left(\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} \right) \neq \frac{1-\beta}{1-\alpha} \left(\frac{(\mu-1)\lambda}{\rho - n + (\sigma-1)g_c + \lambda} \right) = \frac{s_L}{1-s_L},$$

(52) $$\frac{s_K^*}{1-s_K^*}=\frac{\beta}{\alpha}\left(\frac{\gamma g_A}{\rho-n+(\sigma-1)g_c+(1-\phi)g_A}\right)\neq\frac{\beta}{\alpha}\left(\frac{(\mu-1)\lambda}{\rho-n+(\sigma-1)g_c+\lambda}\right)=\frac{s_K}{1-s_K}.$$

(51) and (52) indicate the various sources of externalities and distortion: (a) the negative congestion externality $\gamma\in(0,1]$; (b) the positive or negative externality in intertemporal knowledge spillovers $\phi\in(-\infty,1)$; (c) the static surplus appropriability problem $(\mu-1)/\mu\in(0,1]$, which is a positive externality; (d) the distortion of patent protection in driving a wedge of $\mu>1$ between the factor payment for production inputs and their marginal products; and (e) the positive externality of creative destruction together with the negative externality of the business-stealing effect given by the difference between $\lambda/(\rho-n+(\sigma-1)g_c+\lambda)$ and $g_A/(\rho-n+(\sigma-1)g_c+g_A)$. In addition, in the case of suboptimal profit-sharing arrangements between patentholders, both $s_L/(1-s_L)$ and $s_K/(1-s_K)$ are decreased by the backloading discount factor that is a non-decreasing function of leading breadth.[17] Given the existence of positive and negative externalities, it requires a numerical calibration to the data that will be performed in Section 3 to determine whether the market economy over- or under-invests in R&D.

If the market economy underinvests in R&D as also suggested by Jones and Williams (1998) and (2000), the government can increase patent breadth to reduce the extent of market failures. However, as Propositions 2 demonstrates, an increase in η mitigates the problem of underinvestment in R&D at the cost of worsening the dynamic distortion on capital accumulation. At the constrained social optimum, the government balances these two effects or until patent breadth loses its effectiveness when $z^\eta=1/\varepsilon$.

2.12. Second-Best Optimal Patent Breadth

Given the market equilibrium conditions for $i(\eta)$, $s_L(\eta)$ and $s_K(\eta)$, the benevolent government chooses the second-best optimal level of patent breadth η^* by maximizing

[17] Refer to equation (55).

$$U = \int_0^{\infty} e^{-(\rho-n)t} \frac{c_t^{1-\sigma}(\eta)}{1-\sigma} dt \tag{53}$$

subject to the aggregate production function, the law of motion for capital, and the law of motion for R&D-driven technology. An increase in patent breadth reduces $i(\eta)$ while increases $s_L(\eta)$ and $s_K(\eta)$. If an interior solution such that $z^{\eta^*} \leq 1/\varepsilon$ exists, then the first-order condition that balances the opposing effects on social welfare is given by

$$\omega_I(\eta^*)\frac{\partial i(\eta^*)}{\partial \eta} + \omega_K(\eta^*)\frac{\partial s_K(\eta^*)}{\partial \eta} + \omega_L(\eta^*)\frac{\partial s_L(\eta^*)}{\partial \eta} = 0\,. \tag{*}$$

Each of the ω's represents a weight on the social planner's optimal rule for patent breadth,[18] and the value of each ω is increasing in the difference between the socially optimal and the market-equilibrium levels of its corresponding variable. For example, the further away the market-equilibrium rate of investment is from its social optimum, the larger the weight the social planner should place on $\partial i / \partial \eta$ to prevent patent breadth from increasing the wedge.

Proposition 3: *Suppose there is underinvestment in R&D in the decentralized equilibrium. The first-order condition that characterizes the optimal patent breadth is given by* (*) *if the patent authority enforces the socially optimal profit-sharing arrangement between patentholders.*

3. Calibration

Using the framework developed above, this section provides a quantitative assessment on the effects of eliminating blocking patent and increasing patent breadth. Given the recent policy changes in increasing patent breadth in the 80's, the structural parameters are calibrated using long-run aggregate data of the US's economy from 1953 to 1980. The first numerical exercise considers the effects of eliminating blocking patent on R&D and consumption. The second numerical exercise considers the effects of

[18] See Appendix I for the details. Appendix II derives the analogous expression when the static distortion also exists.

increasing patent breadth to the second-best optimum while holding the effect of blocking patent constant. The results for this exercise are firstly presented for the case in which all sectors are monopolistic and hence the static distortion is absent. Then, the results are presented for the case in which there exist both monopolistic and competitive sectors in order to introduce static distortion into the model and to compare the relative magnitude of the static and dynamic distortionary effects. Finally, the transition dynamics are computed to investigate the effect on consumption during the transitional periods.

3.1. Externality Parameters

The first step is to calibrate the key externality parameters γ (intratemporal duplication) and ϕ (intertemporal spillover). For each value of g_A, g_Z, n, α and β, the balanced-growth condition (45) determines a unique value for $\gamma/(1-\phi)$. The annual average TFP growth rate g_{TFP} is 1.33%,[19] and the labor-force growth rate n is 1.94%.[20] The capital-intensity parameter α in the production sector is set to a conventional value of 0.3, and different plausible values for the R&D capital-intensity parameter $\beta \in \{0, \alpha, 2\alpha, 3\alpha\}$ are considered. $\beta = 0$ corresponds to the knowledge-driven specification in Romer (1990), and $\beta = \alpha$ corresponds to the lab-equipment specification in Rivera-Batiz and Romer (1991) and Jones and Williams (2000). $\beta \in \{2\alpha, 3\alpha\}$ corresponds to the case in which the R&D sector is more capital intensive than the production sector. I will firstly consider the case in which long-run TFP growth is solely driven by R&D (i.e. $g_{TFP} = g_A$ and $g_Z = 0$). Given the above parameters, I firstly calculate the implied value for $\gamma/(1-\phi)$, which is sufficient to determine the new balanced-growth level of consumption. However, holding $\gamma/(1-\phi)$ constant, a larger γ implies a faster convergence rate to the new balanced-growth path; therefore, it is important to consider different values of γ. The calibrated values of ϕ that correspond to a range of values for $\gamma \in [0.1, 1.0]$ are listed in Table 1.

---

[19] Multifactor productivity for the private non-farm business sector is obtained from the Bureau of Labor Statistics.
[20] The data on the annual average size of the labor force is obtained from the Bureau of Labor Statistics.

Table 1: Calibrated Values for φ										
β / γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.85	0.71	0.56	0.42	0.27	0.12	-0.02	-0.17	-0.31	-0.46
α	0.81	0.62	0.43	0.24	0.06	-0.13	-0.32	-0.51	-0.70	-0.89
2α	0.77	0.54	0.30	0.07	-0.16	-0.39	-0.62	-0.85	-1.09	-1.32
3α	0.73	0.45	0.18	-0.10	-0.37	-0.65	-0.92	-1.20	-1.47	-1.75

3.2. First-Best Level of R&D Spending

The second step is to calculate the first-best level of R&D spending, which requires the discount rate, the inverse of the elasticity of intertemporal substitution and the empirical markup. The discount rate is set to a conventional value of 0.04, and the elasticity of intertemporal substitution (i.e. $1/\sigma$) is set to 0.25.[21] The implied real interest rate given by $r = \rho + g_{TFP}\sigma/(1-\alpha)$ is 11.6%, which is higher than the historical real rate of return in the US's stock market, and this higher interest rate implies a lower level of first-best R&D spending. As a result, the model is less likely to overestimate the extent of R&D underinvestment. For the empirical markup μ, I make use of Laitner and Stolyarov's (2004) estimate of 1.10 (i.e. a 10% aggregate markup).[22] Given these additional parameters, I firstly calculate the calibrated values for the following useful ratio $\gamma g_A/(\rho - n + (\sigma - 1)g_c + (1-\phi)g_A)$ that appears in (51) and (52) for a range of values for $\gamma \in [0.1, 1.0]$ and $\beta \in \{0, \alpha, 2\alpha, 3\alpha\}$ in Table 2.

Table 2: Calibrated Values for $\gamma\, g_A/(\rho\text{-}n+(\sigma\text{-}1)g_c+(1\text{-}\phi)g_A)$										
β / γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.02	0.03	0.05	0.06	0.08	0.09	0.10	0.11	0.13	0.14
α	0.02	0.03	0.05	0.06	0.07	0.09	0.10	0.11	0.12	0.13
2α	0.02	0.03	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12
3α	0.02	0.03	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12

[21] It is well-known that there is a discrepancy between the estimated elasticity of intertemporal substitution from dynamic macro models (closed to 1) and econometric studies (closed to 0). Guvenen (2006) shows that this difference is due to the heterogeneity in households' preferences and wealth inequality. In short, the *average investor* has a high elasticity of intertemporal substitution while the *average consumer* has a much lower elasticity. Since my interest is in the effects on consumption, I calibrate the value of σ according to the average consumer.

[22] Basu and Fernald (1997) estimate that the aggregate profit share in the US is about 3%. Assuming cost minimization, the return to scale = markup x (1 - the profit share). Basu and Fernald's (1997) estimates also suggest that "a typical industry has roughly constant returns to scale." (p. 250) I prefer the larger empirical markup from Laitner and Stolyarov (2004) for the following reason. A larger markup leads to more significant distortionary effects and hence is less likely to overestimate the net social benefits of increasing R&D and patent breadth.

(51) and (52) show that the extent of R&D underinvestment is determined by the relative magnitude between $\gamma g_A/(\rho - n + (\sigma - 1)g_c + (1-\phi)g_A)$ and $(\mu - 1)\lambda/(\rho - n + (\sigma - 1)g_c + \lambda)$. Although the calibrated values for ϕ in Table 1 vary across different values of β, the calibrated values for $\gamma g_A/(\rho - n + (\sigma - 1)g_c + (1-\phi)g_A)$ vary only slightly across β. Therefore, I will consider $\beta = \alpha$ that yields convenient analytical expressions as the benchmark. The first-best level of R&D spending as a share of GDP for $\beta = \alpha$ is given in Table 3.

Table 3: First-Best Optimal R&D Shares

γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
R&D	1.6%	3.1%	4.5%	5.7%	6.9%	7.9%	8.9%	9.8%	10.7%	11.5%

The average ratio of private spending on R&D to GDP in the US is 1.15% between 1953 and 1980;[23] therefore, the model predicts that there was a severe degree of underinvestment in R&D before the increase in patent breadth in the 80's. This finding of underinvestment in R&D is consistent with Jones and Williams (2000).

However, Comin (2004) argues that Jones and Williams' (2000) finding is based on the assumption that long-run TFP growth is entirely driven by R&D, which is still an open empirical question. To consider this critique,

(54) $$g_A = \xi g_{TFP},$$

where $\xi \in [0,1]$ captures the fraction of long-run TFP growth that is driven by R&D. The remaining fraction of long-run TFP growth is driven by the exogenous process Z_t such that $g_Z = (1-\xi)g_{TFP}$. Given this modified setting, I plot the first-best R&D shares for $\xi \in [0,1]$ in Figure 3.

[insert Figure 3 here]

Figure 3 shows that there was underinvestment in R&D prior to 1980 over a wide range of parameters. To determine the empirically relevant range for the values of ξ essentially requires answering a much bigger

[23] The data is obtained from the National Science Foundation and the Bureau of Economic Analysis. R&D is net of federal spending, and GDP is net of government spending. The observations in the data series of R&D spending are missing for 1954 and 1955.

question, "what are the factors that drive long-run TFP growth in the data?" This is certainly a very important question but beyond the scope of the current paper. Therefore, I will leave it to the readers to decide on their preferred numbers in Figure 3 and continue presenting results for a range of parameters.

3.3. Blocking Patent

The share of R&D in the data, denoted by R&D/GDP, corresponds to s_r / μ in the model. Because $\alpha = \beta$, $s_r = s_L = s_K$. In the case of suboptimal profit-sharing arrangement, s_r is given by

(55)
$$\frac{s_r}{1-s_r} = \left(\frac{(\mu-1)\lambda}{\rho - n + (\sigma-1)g_c + \lambda} \right) \nu(\eta_{lead}),$$

where $\nu(\eta_{lead})$ is the backloading discount factor, whose value depends on the profit-sharing arrangement between patentholders. In the case of *complete frontloading*, $\nu(\eta_{lead}) = 1$.

Lemma 4: *In the case of the complete backloading profit-sharing arrangement, the backloading discount factor is given by* $\nu(\eta_{lead}) = (\lambda / (r + \lambda - g_Y))^{\eta_{lead}}$.

The steady-state value of ν can be calibrated from the following condition

(56)
$$\nu = \frac{\mu(R\,\&\,D/GDP)}{1-\mu(R\,\&\,D/GDP)} \Bigg/ \left(\frac{(\mu-1)\lambda}{\rho - n + (\sigma-1)g_c + \lambda} \right).$$

The calibrated values of ν for a range of values for $\lambda = [0.04, 0.20]$ are in Table 4, which suggests a severe problem of blocking patent in the economy.

Table 4: Calibrated Values for ν

λ	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
ν	0.376	0.293	0.252	0.227	0.210	0.199	0.190	0.183	0.177

A number of studies has estimated the arrival rate of innovations λ (i.e. the obsolescence probability for a patent). For example, Lanjouw (1998) structurally estimates a patent renewal model using patent renewal data in a number of industries from Germany,[24] and the estimated probability of obsolescence ranges from 7% for computer patents to 12% for engine patents. Also, a conventional value for the rate of depreciation in patent value is about 15% (e.g. Pakes (1986)). In the current model, the patent-value depreciation rate is given by $\lambda - g_Y$, which implies that λ should be at least 15%. On the other hand, Caballero and Jaffe (2002) estimate a mean rate of creative destruction of about 4%. The average empirical estimate for λ of 10% from these studies will be taken as the benchmark.

Upon eliminating blocking patent (i.e. setting $\nu = 1$), s_r and R&D/GDP would increase substantially to the values in Table 5.

Table 5: R&D Shares without Blocking Patent

λ	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
s_r	3.3%	4.2%	4.8%	5.3%	5.7%	6.0%	6.3%	6.5%	6.7%
R&D	3.0%	3.8%	4.4%	4.8%	5.2%	5.5%	5.7%	5.9%	6.1%

In the following, the effect of eliminating blocking patent is expressed in terms of the percentage change in the balanced-growth level of consumption per year. Along the balanced-growth path, per capita consumption increases at an exogenous growth rate g_c. Therefore, after dropping the exogenous growth path and some constant terms and solving for the balanced-growth level of technology and capital-labor ratio, the expression for the balanced-growth level of consumption that depends on the capital investment rate $i(\mu)$ and the R&D share $s_r(\mu)$ simplifies to

(57) $$c_0(\mu) = \left(i(\mu)^{\frac{\alpha(1-\phi+\gamma)}{(1-\alpha)(1-\phi)-\alpha\gamma}} (1-i(\mu))\, s_r(\mu)^{\frac{\gamma}{(1-\alpha)(1-\phi)-\alpha\gamma}} (1-s_r(\mu))^{\frac{(1-\phi)}{(1-\alpha)(1-\phi)-\alpha\gamma}} \right).$$ [25]

Therefore, the percentage change in the long-run consumption can be decomposed into four terms.

[24] The studies in this empirical literature are mostly based on European data. In the US, patent maintenance fees were not initiated until 1982, and the fees are due 3.5 years ($900), 7.5 years ($2300) and 11.5 years ($3800) after a patent is granted, rather than annually as in some European countries.
[25] Refer to Appendix II for the derivation of the corresponding expression for the general case with static distortion.

$$(58)\quad \Delta \ln c_0(\mu) = \begin{pmatrix} \left(\dfrac{\alpha(1-\phi+\gamma)}{(1-\alpha)(1-\phi)-\alpha\gamma}\right)\Delta \ln i(\mu) + \Delta \ln(1-i(\mu)) + \\ \left(\dfrac{\gamma}{(1-\alpha)(1-\phi)-\alpha\gamma}\right)\Delta \ln s_r(\mu) + \left(\dfrac{(1-\phi)}{(1-\alpha)(1-\phi)-\alpha\gamma}\right)\Delta \ln(1-s_r(\mu)) \end{pmatrix}.$$[26]

Figure 4 shows that eliminating blocking patent should have a substantial positive effect on long-run consumption unless ξ is very small. Also, a back-of-the-envelope calculation shows that the change in consumption mostly comes from $(\gamma/((1-\alpha)(1-\phi)-\alpha\gamma))\Delta \ln s_r(\mu)$; in other words, other general-equilibrium effects only have secondary impacts on long-run consumption.

[insert Figure 4 here]

3.4. Optimal Patent Breadth

The next numerical exercise computes the second-best optimal markup μ^*. If the empirical markup is below μ^*, then, increasing patent breadth in order to stimulate R&D would improve social welfare.[27] Otherwise, patent breadth should be reduced. In the followings, a range of values for $\varepsilon \in [0, 0.9]$ will be considered, and each value of ε corresponds to a unique value of z according to $\tilde{z}(\varepsilon) = g_A / \lambda$.

Table 6: The Calibrated Values for z

ξ / ε	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.2	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.02
0.4	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.04
0.6	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.07	1.06
0.8	1.11	1.11	1.11	1.11	1.11	1.11	1.10	1.10	1.09	1.08
1.0	1.14	1.14	1.14	1.14	1.14	1.13	1.13	1.12	1.11	1.09

Given an empirical markup of 1.1, Table 6 indicates non-zero leading breadth for a wide range of parameters (i.e. when $z < 1.1$), and this finding is consistent with the backloading discount factor being

[26] Note that the coefficients in (58) are solely determined by α and $\gamma/(1-\phi)$.

[27] Social welfare refers to the lifetime utility of the representative household.

less than 1. Using Broda and Weinstein's (2006) empirical estimates,[28] I will set $\varepsilon = 0.8$ but also consider $\varepsilon = 0$ that corresponds to the standard Cobb-Douglas specification.

Figures 5a and 5b present the second-best optimal markup given by $\mu^* \equiv \min\{z^{\eta^*}, 1/\varepsilon\}$ that can be computed by numerically solving (*). The calculation assumes that patent breadth increases while the effect of blocking patent (i.e. the backloading discount factor ν) remains constant.[29]

[insert Figure 5a here]

For $\varepsilon = 0.8$, the constraint $\mu^* \equiv \min\{z^{\eta^*}, 1/\varepsilon\}$ becomes binding for certain parameters. Also, there is insufficient patent protection (i.e. the empirical markup is too small) for a wide range of parameters.

[insert Figure 5b here]

For $\varepsilon = 0$, the optimal markup is simply given by z^{η^*}. As before, there is insufficient patent protection for a wide range of parameters. In addition, the optimal levels of markup are almost identical for $\varepsilon = \{0, 0.8\}$ unless the constraint $\mu^* \equiv \min\{z^{\eta^*}, 1/\varepsilon\}$ becomes binding.

Finally, the effects of changing the empirical markup from 1.1 to the second-best optimum are expressed in terms of the percentage change in the balanced-growth level of consumption per year. Figure 6a presents the percentage change in long-run consumption for $\varepsilon = 0.8$.

[insert Figure 6a here]

For the range of parameters that involves insufficient patent breadth, increasing patent breadth to the optimal level can lead to a substantial increase in long-run consumption. For the range of parameters that involves excessive patent breadth (i.e. low values of γ and ξ), reducing the markup to the optimal level leads to a fall in long-run consumption due to the fall in s_r but improves welfare.[30] In the case of $\xi = 0$ (i.e. when R&D is completely wasteful), the increase in long-run consumption is driven by the

[28] See footnote 9.
[29] Refer to Appendix II for a similar derivation of introducing a constant ν into the optimal rule of patent breadth.
[30] Note that maximizing long-run consumption is not the same as maximizing social welfare.

reallocation of resources to the production sector. Figure 6b presents the percentage change in consumption for $\varepsilon = 0$ and shows very similar results.

[insert Figure 6b here]

3.5. Dynamic vs. Static Distortions

As mentioned before, the assumption that all intermediate-goods industries are monopolistic eliminates the static distortion. As a result, the previous numerical exercise overestimates the effect of increasing patent breadth on consumption. The assumption is made in order to maintain the analytical tractability of the aggregate conditions under the CES aggregator. In this subsection, this assumption is relaxed under the special case of a Cobb-Douglas aggregator. The followings sketch out the key equations, and the details of the changes in the model are provided in Appendix II.

The production function for the final goods becomes

$$Y_t = \exp\left(\int_0^1 \ln X_t(j)dj \right). \tag{59}$$

Among the continuum of the intermediate goods $j \in [0,1]$, a fraction θ of the industries is characterized by perfect competition. Without loss of generality, the industries are ordered such that $P_t(j') = MC_t(j')$ for $j' \in [0, \theta]$. The aggregate price level is

$$P_t = \mu MC_t, \tag{60}$$

where $\mu \equiv (z^{\eta})^{1-\theta}$ is the average markup in the economy, and z^{η} is the markup in monopolistic industries. Because of the markup pricing in monopolistic industries and the marginal-cost pricing in competitive industries, the industries' ratio of factor inputs is z^{η}. Using this information, the aggregate production function becomes

$$Y_t = \vartheta(\eta) A_t Z_t K_{x,t}^{\alpha} L_{x,t}^{1-\alpha}. \tag{61}$$

For $\theta \in (0,1)$, $\vartheta(\eta) \in (0,1)$ represents the static distortionary effect of markup pricing and is defined as

(62) $$\vartheta(\eta) \equiv \frac{(z^{\eta})^{\theta}}{z^{\eta}\theta + (1-\theta)} \in (0,1).$$

Markup pricing in the monopolistic industries distorts production towards the competitive industries and reduces the output of the final goods. Also, $\vartheta(\eta)$ is initially decreasing in θ and subsequently increasing with $\vartheta(\eta) = 1$ for $\theta = \{0,1\}$. Therefore, at least over a range of parameters, the static distortionary effect is increasing in the fraction of competitive industries.

Given this setup, I will once again numerically evaluate the change in the balanced-growth level of consumption from changing the level of patent breadth to the constrained optimum. R&D/GDP in the data relates to s_r in the model according to

(63) $$\frac{R\,\&\,D}{GDP} = s_r \left(\frac{z^{\eta}\theta + (1-\theta)}{z^{\eta}} \right).$$

An additional parameter that is needed for this exercise is the fraction of competitive industries θ, and results are provided for $\theta \in \{0.25, 0.5\}$. With the static distortionary effect and a constant backloading discount factor, the first-order condition that characterizes the optimal patent breadth is given by

(**) $$\omega_{\theta}(\eta)\frac{\partial \vartheta(\eta)}{\partial \eta} + \omega_{I}(\eta)\frac{\partial i(\eta)}{\partial \eta} + \omega_{K}(\eta)\frac{\partial s_K(\eta)}{\partial \eta} + \omega_{L}(\eta)\frac{\partial s_L(\eta)}{\partial \eta} = 0.$$ [31]

Figure 7a provides the second-best optimal level of average markup μ^* for $\theta = 0.25$.

[insert Figure 7a here]

Comparing with Figure 5b (i.e. without the static distortion), the optimal level of markup is now smaller than before because of the static distortionary effect. However, the qualitative results remain unchanged that there is insufficient patent breadth over a wide range of parameters. Figure 7b provides the second-best optimal level of average markup μ^* for $\theta = 0.5$. In the case, the optimal markups become even smaller because the static distortionary effect are more severe at a higher value of θ. In summary, there is still insufficient patent breadth over a wide range of parameters.

[31] See Appendix II for the derivation.

[insert Figure 7b here]

Finally, the effects of changing the empirical markup from 1.1 to the second-best optimum are expressed in terms of the percentage change in the balanced-growth level of consumption per year. With static distortion, the expression for the balanced-growth level of consumption becomes

$$(64) \qquad c_0(\eta) = \left(\tilde{\vartheta}(\eta)^{1/(1-\alpha)}\, i(\eta)^{\frac{\alpha(1-\phi+\gamma)}{(1-\alpha)(1-\phi)-\alpha\gamma}} (1-i(\eta))\, s_r(\eta)^{\frac{\gamma}{(1-\alpha)(1-\phi)-\alpha\gamma}} (1-s_r(\eta))^{\frac{(1-\phi)}{(1-\alpha)(1-\phi)-\alpha\gamma}} \right).$$ [32]

Figure 8a and 8b provide the percentage change in long-run consumption for $\theta \in \{0.25, 0.5\}$ assuming the backloading discount factor to be constant as before.

[insert Figures 8a and 8b here]

Two important results emerge. Firstly, the positive effect on consumption is substantial unless either ξ or γ is very small. Therefore, taking the static distortionary effect into consideration does not alter the previous finding that increasing patent breadth could mitigate the R&D underinvestment problem and increase consumption. Secondly, the magnitude of the increase in consumption becomes smaller as θ increases because the static distortionary effect of markup pricing becomes more severe. Tables 7 and 8 provide the partial effects on consumption from the static distortion given by $(1-\alpha)^{-1}\Delta \ln \tilde{\vartheta}(\eta)$, and the dynamic distortion given by $\left(\dfrac{\alpha(1-\phi+\gamma)}{(1-\alpha)(1-\phi)-\alpha\gamma}\right)\Delta \ln i(\eta) + \Delta \ln(1-i(\eta))$. The following tables show that as θ increases, the static distortionary effect becomes more significant but is still smaller than the dynamic distortionary effect. Furthermore, both effects are relatively insignificant compared to the change in consumption.

[32] See Appendix II for the derivation.

Table 7a: Static Distortion for θ = 0.25					
ξ / γ	0.20	0.40	0.60	0.80	1.00
0.0	0.2%	0.2%	0.2%	0.2%	0.2%
0.2	0.2%	0.1%	0.0%	-0.1%	-0.2%
0.4	0.1%	-0.1%	-0.4%	-0.6%	-0.9%
0.5	0.0%	-0.3%	-0.6%	-1.0%	-1.3%
0.6	0.0%	-0.4%	-0.9%	-1.3%	-1.7%
0.8	-0.2%	-0.8%	-1.4%	-2.0%	-2.6%
1.0	-0.3%	-1.2%	-2.0%	-2.8%	-3.5%

Table 8a: Dynamic Distortion for θ = 0.25					
ξ / γ	0.20	0.40	0.60	0.80	1.00
0.0	1.8%	1.8%	1.8%	1.8%	1.8%
0.2	1.3%	0.6%	0.0%	-0.4%	-0.8%
0.4	0.6%	-0.8%	-1.9%	-2.8%	-3.5%
0.5	0.2%	-1.7%	-3.1%	-4.2%	-5.1%
0.6	-0.3%	-2.6%	-4.3%	-5.6%	-6.7%
0.8	-1.5%	-4.7%	-7.1%	-8.9%	-10.4%
1.0	-3.0%	-7.3%	-10.4%	-12.8%	-14.7%

Table 7b: Static Distortion for θ = 0.5					
ξ / γ	0.20	0.40	0.60	0.80	1.00
0.0	0.6%	0.6%	0.6%	0.6%	0.6%
0.2	0.6%	0.4%	0.2%	0.1%	-0.1%
0.4	0.4%	0.0%	-0.4%	-0.8%	-1.2%
0.5	0.3%	-0.3%	-0.8%	-1.3%	-1.7%
0.6	0.2%	-0.5%	-1.1%	-1.7%	-2.3%
0.8	-0.1%	-1.0%	-1.9%	-2.7%	-3.4%
1.0	-0.3%	-1.5%	-2.7%	-3.7%	-4.6%

Table 8b: Dynamic Distortion for θ = 0.5					
ξ / γ	0.20	0.40	0.60	0.80	1.00
0.0	1.7%	1.7%	1.7%	1.7%	1.7%
0.2	1.4%	0.8%	0.4%	0.1%	-0.1%
0.4	1.0%	-0.1%	-0.8%	-1.3%	-1.8%
0.5	0.7%	-0.6%	-1.4%	-2.1%	-2.7%
0.6	0.4%	-1.1%	-2.2%	-3.0%	-3.6%
0.8	-0.2%	-2.4%	-3.8%	-4.9%	-5.7%
1.0	-1.1%	-3.9%	-5.7%	-7.1%	-8.1%

3.6. Transition Dynamics

The purpose of this subsection is to compute the entire growth path of per capita consumption after the broadening of patent breadth. Again, this exercise is performed for the special case of the Cobb-Douglas aggregator, and all industries are assumed to be monopolistic. The dynamics of the model is characterized by the following four differential equations. The capital stock is a predetermined variable and evolves according to

(65) $$\dot{K}_t = Y_t - C_t - K_t\delta .$$

The aggregate technology is also a predetermined variable and evolves according to

(66) $$\dot{A}_t = A_t\lambda_t \ln z .$$[33]

Consumption is a jump variable and evolves according to the Euler equation

---

[33] This convenient expression is derived as $\ln A_t = \left(\int_0^1 \ln z^{m_t(j)} dj\right) = \left(\int_0^1 m_t(j)dj\right)\ln z = \left(\int_0^t \lambda(\tau)d\tau\right)\ln z$; then, simple differentiation yields $\dot{A}_t / A_t = \lambda_t \ln z$.

(67) $$\dot{c}_t = c_t(r_t - \rho)/\sigma .$$

The aggregate value of patents is a also a jump variable and evolves according to

(68) $$\dot{V}_t = (r_t + \lambda_t)V_t - \nu\left(\frac{\mu - 1}{\mu}\right)Y_t ,$$

where the backloading discount factor ν is assumed to be constant and equal its steady-state value.[34]

At the aggregate level, the generalized quality-ladder model is very similar to the model in Jones (1995b), whose dynamic properties have been investigated by a number of recent studies. For example, Arnold (2006) analytically derives the uniqueness and stability of the steady state with certain parameter restrictions. Steger (2005) and Trimborn, Koch and Steger (2006) numerically evaluate the transition dynamics of the model. In summary, to solve the model, I firstly transform $\{c_t, V_t, K_t, A_t\}$ in the four differential equations into its stationary form,[35] and then, compute the transition dynamics from the old steady state to the new one using the relaxation algorithm developed by Trimborn *et al* (2006).

Figure 9a compares the transition path (in blue) of log consumption per capita with its original balanced-growth path (in red) and its new balanced-growth path (in green) for the following parameters: $\xi = 0.5$ and $\gamma = 0.55$. In this case, the optimal markup is 1.25, and long run consumption increases by 31%.

[insert Figures 9a here]

Upon the strengthening of patent protection, consumption per capita gradually rises towards the new balanced growth path. Although factor inputs shift towards the R&D sector and the output of final goods drops as a result, the possibility of investing less and running down the capital stock enables consumption smoothing. To compare with previous studies, such as Kwan and Lai (2003), Figure 9b presents the transition dynamics for $\delta = 1$ (i.e. complete capital depreciation). In this case, the result is consistent with Kwan and Lai (2003) that consumption falls in response to the strengthening of patent protection.

---

[34] Although the variation in the arrival rate of innovations may cause the backloading discount factor to vary along the transition path, its value is very difficult to determine. Therefore, a simple approximation is made here.
[35] Refer to Appendix III for the details.

[insert Figures 9b here]

To ensure the robustness of this finding, a sensitivity analysis has been performed for different values of ξ and γ. At a larger value of either ξ and γ, consumption increases by even more on impact. A larger ξ also implies a higher position of the new balanced-growth path. Holding ξ constant, a larger γ implies a faster rate of convergence. When both ξ and γ are smaller than 0.45, the household suffers small consumption losses during the initial phase of the transition path. However, when either ξ or γ is closed to one, the other parameter could be as small as 0.25 without causing short-run consumption losses. In summary, strengthening patent protection does not always lead to short-run consumption losses.

4. Conclusion

This paper has attempted to accomplish three objectives. The first objective is to develop a tractable framework for a dynamic general-equilibrium analysis on optimal patent breadth. The second objective is to analyze the dynamic distortion on capital accumulation that has been neglected by previous studies on patent policy. The third objective is to provide a quantitative assessment on the effects of eliminating blocking patent and increasing patent breadth. The calibration exercise suggests a number of findings. Firstly, the market economy underinvests in R&D so long as a non-negligible fraction of long-run TFP growth is driven by R&D. Secondly, increasing patent breadth may be an effective solution to this potential problem of R&D underinvestment, and the resulting effect on consumption can be substantial.

However, the readers should interpret the numerical results with some important caveats in mind. The first obvious caveat is that although the quality-ladder model has been generalized as an attempt to capture more realistic features of the economy, it is still an oversimplification of the real world. In particular, the finding of patent policy having a substantial positive effect on consumption is based on the assumptions that a non-negligible fraction of long-run TFP growth is driven by R&D and the incentive for private investments in R&D increases in response to broadening patent protection. The validity of these assumptions remains as an empirical question. Therefore, the numerical results should be viewed as illustrative at best. The second caveat is that the representative-agent setting ignores the distributional consequences of increasing patent protection, and the efficiency-equity tradeoff should be carefully considered by policymakers. The third caveat is that the model is based on a closed-economy setting. If the suboptimal level of patent protection arises from a multi-country Nash equilibrium, a unilateral deviation from its social best response function would render a country worse off despite the increase in R&D. In this case, the numerical results should be interpreted as the effects of increasing patent protection from the Nash equilibrium to a more cooperative symmetric equilibrium. The Nash equilibrium is globally suboptimal because of the detrimental effects of international free-riding on innovations.[36]

[36] See, e.g. Grossman and Lai (2004) for an elegant formulation of this insight.

References

1. **Aghion, Phillippe; and Howitt, Peter** (1992) "A Model of Growth through Creative Destruction" *Econometrica* vol. 60, p. 323-351.
2. **Arnold, Lutz G.** (2006) "The Dynamics of the Jones R&D Growth Model" *Review of Economic Dynamics* vol. 9, p.143-152.
3. **Barro, Robert J.; and Sala-i-Martin, Xavier** (2003) "Economic Growth" The MIT Press.
4. **Basu, Susanto; and Fernald, John G.** (1997) "Returns to Scale in U.S. Production: Estimates and Implications" *Journal of Political Economy* vol. 105, p. 249-283.
5. **Broda, Christian; and Weinstein, David E.** (2006) "Globalization and the Gains from Variety" *Quarterly Journal of Economics* vol. 121, p. 541-585.
6. **Caballero, Ricardo J.; and Jaffe, Adam B.** (2002) "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth" in A.B. Jaffe and M. Trajtenberg, eds., *Patents, Citations and Innovations: A Window on the Knowledge Economy* p. 89-152.
7. **Chu, Angus C.** (2007) "Optimal Patent Length: Quantifying the Effects of Patent Extension" *University of Michigan* Working Paper.
8. **Comin, Diego** (2004) "R&D: A Small Contribution to Productivity Growth" *Journal of Economic Growth* vol. 9, p. 391-421.
9. **Futagami, Koichi; and Iwaisako, Tatsuro** (2007) "Dynamic Analysis of Patent Policy in an Endogenous Growth Model" *Journal of Economic Theory* vol. 132, p. 306-334.
10. **Gallini, Nancy T.** (2002) "The Economics of Patents: Lessons from Recent U.S. Patent Reform" *Journal of Economic Perspectives* vol. 16, p. 131-154.
11. **Gilbert, Richard; and Shapiro, Carl** (1990) "Optimal Patent Length and Breadth" *RAND Journal of Economics* vol. 21, p. 106-112.
12. **Goh, Ai-Ting; and Olivier, Jacques** (2002) "Optimal Patent Protection in a Two-Sector Economy" *International Economic Review* vol. 43, p. 1191-1214.

13. **Green, Jerry R.; and Scotchmer, Suzanne** (1995) "On the Division of Profit in Sequential Innovation" *RAND Journal of Economics* vol. 26, p. 20-33.

14. **Grossman, Gene M.; and Helpman, Elhanan** (1991) "Quality Ladders in the Theory of Growth" *Review of Economic Studies* vol. 58, p. 43-61.

15. **Grossman, Gene M.; and Lai, Edwin L.-C.** (2004) "International Protection of Intellectual Property" *American Economic Review* vol. 94, p. 1635-1653.

16. **Guvenen, Fatih** (2006) "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective" *Journal of Monetary Economics* vol. 53, p. 1451-1472.

17. **Hall, Bronwyn H.; Jaffe, Adam B.; and Trajtenberg, Manuel** (2002) "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools" in A.B. Jaffe and M. Trajtenberg, eds., *Patents, Citations and Innovations: A Window on the Knowledge Economy* p. 403-459.

18. **Helpman, Elhanan** (1993) "Innovation, Imitation and Intellectual Property Rights" Econometrica vol. 61, p. 1247-1280.

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

20. **Jaffe, Adam B.** (2000) "The U.S. Patent System in Transition: Policy Innovation and the Innovation Process" *Research Policy* vol. 29, p. 531-557.

21. **Jaffe, Adam B.; and Lerner, Josh** (2004) "Innovation and Its Discontents: How Our Broken System Is Endangering Innovation and Progress, and What to Do About It" Princeton, NJ: Princeton University Press.

22. **Jones, Charles I.** (1995a) "Time Series Tests of Endogenous Growth Models" *Quarterly Journal of Economics* vol. 110, p. 495-525.

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

24. **Jones, Charles I.** (1999) "Growth: With or Without Scale Effects" *American Economic Review Papers and Proceedings* vol. 89 p. 139-144.

25. **Jones, Charles I.; and Williams, John C.** (1998) "Measuring the Social Return to R&D" *Quarterly Journal of Economics* vol. 113, p. 1119-1135.

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

27. **Judd, Kenneth L.** (1985) "On the Performance of Patents" *Econometrica* vol. 53, p.567-586.

28. **Klemperer, Paul** (1990) "How Broad Should the Scope of Patent Protection Be? *RAND Journal of Economics* vol. 21, p. 113-130.

29. **Kortum, Samuel; and Lerner, Josh** (1998) "Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?" *Carnegie-Rochester Conference Series on Public Policy* vol. 48, p. 247-304.

30. **Kwan, Yum K.; and Lai, Edwin L.-C.** (2003) "Intellectual Property Rights Protection and Endogenous Economic Growth" *Journal of Economic Dynamics and Control* vol. 27, p. 853-873.

31. **Laitner, John** (1982) "Monopoly and Long-Run Capital Accumulation" *Bell Journal of Economics* vol. 13, p. 143-157.

32. **Laitner, John; and Stolyarov, Dmitriy** (2004) "Aggregate Returns to Scale and Embodied Technical Change: Theory and Measurement Using Stock Market Data" *Journal of Monetary Economics* vol. 51, p. 191-233.

33. **Laitner, John; and Stolyarov, Dmitriy** (2005) "Owned Ideas and the Stock Market" *University of Michigan* Working Paper.

34. **Lanjouw, Jean Olson** (1998) "Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value" *Review of Economic Studies* vol. 65, p. 671-710.

35. **Lerner, Josh** (1994) "The Importance of Patent Scope: An Empirical Analysis" *RAND Journal of Economics* vol. 25, p. 319-333.

36. **Li, Chol-Won** (2001) "On the Policy Implications of Endogenous Technological Progress" *Economic Journal* vol. 111, p. 164-179.

37. **Nordhaus, William** (1969) "Invention, Growth, and Welfare" Cambridge, Mass.: MIT Press.

38. **O'Donoghue, Ted** (1998) "A Patentability Requirement for Sequential Innovation" *RAND Journal of Economics* vol. 29, p. 654-679.

39. **O'Donoghue, Ted; Scotchmer, Suzanne; and Thisse, Jacques-Francois** (1998) "Patent Breadth, Patent Life, and the Pace of Technological Progress" *Journal of Economics and Management Strategy* vol. 7, p. 1-32.

40. **O'Donoghue, Ted; and Zweimuller, Josef** (2004) "Patents in a Model of Endogenous Growth" *Journal of Economic Growth* vol. 9, p. 81-123.

41. **Pakes, Ariel** (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks" *Econometrica* vol. 54, p. 755-784.

42. **Rivera-Batiz, Luis A.; and Romer, Paul M.** (1991) "Economic Integration and Endogenous Growth" *Quarterly Journal of Economics* vol. 106, p. 531-555.

43. **Romer, Paul M.** (1990) "Endogenous Technological Change" *Journal of Political Economy* vol. 98, S71-S102.

44. **Scotchmer, Suzanne** (2004) "Innovation and Incentives" Cambridge, Mass.: MIT Press.

45. **Segerstrom, Paul S.** (1998) "Endogenous Growth without Scale Effects" *American Economic Review* vol. 88, p. 1290-1310.

46. **Steger, Thomas M.** (2005) "Non-Scale Models of R&D-based Growth: The Market Solution" *Topics in Macroeconomics* vol. 5 (1), Article 3.

47. **Stokey, Nancy L.** (1995) "R&D and Economic Growth" *Review of Economic Studies* vol. 62, p.469-489.

48. **Tandon, Pankaj** (1982) "Optimal Patents with Compulsory Licensing" *Journal of Political Economy* vol. 90, p. 470-486.

49. **Trimborn, Timo; Koch, Karl-Josef; and Steger, Thomas M.** (2006) "Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure" *CESifo* Working Paper Series No. 1745.

Appendix I

Proposition 1a: *For any given level of patent breadth, the complete frontloading profit-sharing arrangement is socially optimal if there is underinvestment in R&D in the decentralized equilibrium.*

Proof: Any profit-sharing arrangement that involves the backloading of payoffs reduces the present value of the stream of profits generated by a patent. Denote $\nu \in (0,1)$ to capture this backloading effect

$$V_t(\nu) = \nu\left(\frac{\mu-1}{\mu}\right)\frac{Y_t}{r+\lambda-g_Y}.$$

The first-order conditions from the R&D sector become

$$\frac{s_L(\nu)}{1-s_L(\nu)} = \frac{1-\beta}{1-\alpha}\left(\frac{\nu(\mu-1)\lambda}{r+\lambda-g_Y}\right),$$

$$\frac{s_K(\nu)}{1-s_K(\nu)} = \frac{\beta}{\alpha}\left(\frac{\nu(\mu-1)\lambda}{r+\lambda-g_Y}\right).$$

Note that $i(\nu)$ is also a function of ν because of $s_K(\nu)$ and is given by

$$i(\nu) = \frac{\alpha}{\mu(1-s_K(\nu))}\left(\frac{g_K+\delta}{r+\delta}\right).$$

An rise in ν increases $i(\nu)$, $s_K(\nu)$ and $s_L(\nu)$ and moves them towards to the constrained social optimum if there is underinvestment in R&D. For a given η and μ, determining the socially optimal level of $\nu \in (0,1]$ requires maximizing

$$U = \int_0^\infty e^{-(\rho-n)t}\frac{c_t^{1-\sigma}(\nu)}{1-\sigma}dt$$

subject to the aggregate production function, the law of motion for capital, and the law of motion for technology. The first-order condition for the optimal ν^* is

$$\omega_I(\nu^*)\frac{\partial i(\nu^*)}{\partial \nu} + \omega_K(\nu^*)\frac{\partial s_K(\nu^*)}{\partial \nu} + \omega_L(\nu^*)\frac{\partial s_L(\nu^*)}{\partial \nu} \geq 0.$$

The inequality is strict at the corner solution $\nu^* = 1$. Note that each of the derivatives is strictly positive, and the ω's are defined as

$$\omega_I(\nu) \equiv \left(\begin{array}{l} \left(\alpha + \beta \dfrac{\gamma g_A}{\rho - n + (\sigma - 1) g_c + (1-\phi) g_A} \right) \dfrac{g_K + \delta}{\rho + g_c \sigma + \delta} \\ - \dfrac{i(\nu)}{1 - i(\nu)} \left(1 - \left(\alpha + \beta \dfrac{\gamma g_A}{\rho - n + (\sigma - 1) g_c + (1-\phi) g_A} \right) \dfrac{g_K + \delta}{\rho + g_c \sigma + \delta} \right) \end{array} \right) \frac{1}{i(\nu)},$$

$$\omega_K(\nu) \equiv \left(\beta \frac{\gamma g_A}{\rho - n + (\sigma - 1) g_c + (1-\phi) g_A} - \alpha \frac{s_K(\nu)}{1 - s_K(\nu)} \right) \frac{1}{s_K(\nu)},$$

$$\omega_L(\nu) \equiv \left((1-\beta) \frac{\gamma g_A}{\rho - n + (\sigma - 1) g_c + (1-\phi) g_A} - (1-\alpha) \frac{s_L(\nu)}{1 - s_L(\nu)} \right) \frac{1}{s_L(\nu)}.$$

The outermost bracket in $\omega_I(\nu)$ is strictly positive for $\nu \in (0,1]$ if there is underinvestment in R&D (see Proposition 2). Each of the outermost brackets in $\omega_K(\nu)$ and $\omega_L(\nu)$ is also strictly positive for $\nu \in (0,1]$ if and only if there is underinvestment in R&D. Therefore, the underinvestment in R&D is a *sufficient* condition for $\nu = 1$.■

Proposition 1b: *In the special case of labor being the only factor input for R&D, the complete frontloading profit-sharing arrangement is socially optimal if and only if there is underinvestment in R&D in the decentralized equilibrium.*

Proof: In the special case of labor being the only factor input for R&D, $s_K = 0$ so that i is no longer a function of ν. Consequently, the first-order condition for the optimal ν^* simplifies to

$$\left((1-\beta) \frac{\gamma g_A}{\rho - n + (\sigma - 1) g_c + (1-\phi) g_A} - (1-\alpha) \frac{s_L(\nu^*)}{1 - s_L(\nu^*)} \right) \frac{1}{s_L(\nu^*)} \frac{\partial s_L(\nu^*)}{\partial \nu} \geq 0.$$

The term in the bracket is strictly positive for $\nu \in (0,1]$ *if and only if* there is underinvestment in R&D.■

Proposition 2a: *The decentralized equilibrium rate of investment is below the socially optimal investment rate if either there is underinvestment in R&D or labor is the only factor input for R&D.*

Proof: The socially optimal investment rate i^* is

$$i^* = \left(\alpha + \beta \frac{\gamma g_A}{\rho - n + (\sigma - 1)g_c + (1-\phi)g_A}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta}.$$

The market equilibrium rate of investment i is

$$i = \frac{\alpha}{z^\eta(1-s_K)}\left(\frac{g_K + \delta}{\rho + g_c\sigma + \delta}\right) = \frac{1}{z^\eta}\left(\alpha + \beta \frac{(z^\eta - 1)\lambda}{\rho - n + (\sigma - 1)g_c + \lambda}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta}.$$

Therefore, either $\beta = 0$ or the underinvestment in R&D such that $\frac{\gamma g_A}{\rho - n + (\sigma - 1)g_c + (1-\phi)g_A} >$ $\frac{(z^\eta - 1)\lambda}{\rho - n + (\sigma - 1)g_c + \lambda}$ is *sufficient* for $i^* > i$ because of the markup $z^\eta > 1$.■

Proposition 2b: *An increase in patent breadth leads to a reduction in the decentralized equilibrium rate of investment if the intermediate-goods sector is at least as capital intensive as the R&D sector.*

Proof: Differentiating i with respect to η yields

$$\frac{\partial i}{\partial \eta} = -\frac{\ln z}{z^\eta}\left(\alpha - \beta \frac{\lambda}{\rho - n + (\sigma - 1)g_c + \lambda}\right)\frac{g_K + \delta}{\rho + g_c + \delta}.$$

Since $\rho > n$ by (A1) and $\sigma \geq 1$, $\alpha \geq \beta$ is a *sufficient* condition for $\partial i / \partial \eta < 0$.■

Proposition 3: *Suppose there is underinvestment in R&D in the decentralized equilibrium. The first-order condition that characterizes the optimal patent breadth is given by* (*) *if the patent authority enforces the socially optimal profit-sharing arrangement between patentholders.*

(*) $$\omega_I(\eta)\frac{\partial i(\eta)}{\partial \eta} + \omega_K(\eta)\frac{\partial s_K(\eta)}{\partial \eta} + \omega_L(\eta)\frac{\partial s_L(\eta)}{\partial \eta} = 0.$$

Proof: The second-best optimal level of η can be found by maximizing

(a1) $$U = \int_0^{\infty} e^{-(\rho-n)t} \frac{c_t^{1-\sigma}(\eta)}{1-\sigma} dt$$

subject to the aggregate production function given by

(a2) $$Y_t = A_t Z_t (1-s_K)^{\alpha} (1-s_L)^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha},$$

the law of motion for capital given by

(a3) $$\dot{K}_t = iY_t - K_t\delta,$$

and the law of motion for R&D-driven technology given by

(a4) $$\dot{A}_t = A_t^{\phi}(s_K)^{\beta\gamma}(s_L)^{(1-\beta)\gamma} K_t^{\beta\gamma} L_t^{(1-\beta)\gamma} \varphi\tilde{z}(\varepsilon).$$

The current-value Hamiltonian H is

(a5) $$\begin{aligned} H &= (1-\sigma)^{-1}\left(\frac{(1-i)A_t Z_t (1-s_K)^{\alpha}(1-s_L)^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha}}{L_t}\right)^{1-\sigma} \\ &+ \upsilon_K (iA_t Z_t (1-s_K)^{\alpha}(1-s_L)^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha} - K_t\delta) \\ &+ \upsilon_A A_t^{\phi}(s_K)^{\beta\gamma}(s_L)^{(1-\beta)\gamma} K_t^{\beta\gamma} L_t^{(1-\beta)\gamma} \varphi\tilde{z}(\varepsilon) \end{aligned}.$$

Note that $i(\eta)$, $s_K(\eta)$ and $s_L(\eta)$ are all functions of η. The first-order conditions are

(a6) $$H_K = \frac{\alpha}{K_t}\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma} + \upsilon_K\left(\alpha\frac{i(\eta)Y_t}{K_t} - \delta\right) + \upsilon_A\left(\beta\gamma\frac{\dot{A}_t}{K_t}\right) = (\rho-n)\upsilon_K - \dot{\upsilon}_K,$$

(a7) $$H_A = \frac{1}{A_t}\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma} + \upsilon_K\left(\frac{i(\eta)Y_t}{A_t}\right) + \upsilon_A\left(\phi\frac{\dot{A}_t}{A_t}\right) = (\rho-n)\upsilon_A - \dot{\upsilon}_A,$$

(a8) $$\begin{aligned} H_{\eta} &= -\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma}\left(\frac{1}{1-i(\eta)}\frac{\partial i(\eta)}{\partial\eta} + \frac{\alpha}{1-s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta} + \frac{1-\alpha}{1-s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}\right) \\ &+ \upsilon_K Y_t\left(\frac{\partial i(\eta)}{\partial\eta} - \frac{\alpha}{1-s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta}i(\eta) - \frac{1-\alpha}{1-s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}i(\eta)\right) \\ &+ \upsilon_A\dot{A}_t\left(\frac{\beta\gamma}{s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta} + \frac{(1-\beta)\gamma}{s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}\right) = 0 \end{aligned}.$$

(a6) and (a7) simplify to the following conditions

(a9) $$\upsilon_K K_t(\rho + g_c\sigma + \delta - \alpha(g_K + \delta)) = \alpha((1-i(\eta))Y_t / L_t)^{1-\sigma} + \upsilon_A A_t g_A \beta\gamma,$$

(a10) $$\upsilon_A A_t(\rho - n + (\sigma-1)g_c + (1-\phi)g_A) = ((1-i(\eta))Y_t / L_t)^{1-\sigma} + \upsilon_K K_t(g_K + \delta).$$

Using (a9) and (a10), the first-order condition (a8) that characterizes the second-best optimal level of patent breadth simplifies to

(a11) $$\omega_I(\eta)\frac{\partial i(\eta)}{\partial \eta} + \omega_K(\eta)\frac{\partial s_K(\eta)}{\partial \eta} + \omega_L(\eta)\frac{\partial s_L(\eta)}{\partial \eta} = 0.$$

The ω's are defined as

(a12) $$\omega_I(\eta) \equiv \left(\begin{array}{l} \left(\alpha + \beta\dfrac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A}\right)\dfrac{g_K + \delta}{\rho + g_c\sigma + \delta} \\ -\dfrac{i(\eta)}{1-i(\eta)}\left(1 - \left(\alpha + \beta\dfrac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A}\right)\dfrac{g_K + \delta}{\rho + g_c\sigma + \delta}\right) \end{array} \right)\frac{1}{i(\eta)},$$

(a13) $$\omega_K(\eta) \equiv \left(\beta\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} - \alpha\frac{s_K(\eta)}{1-s_K(\eta)}\right)\frac{1}{s_K(\eta)},$$

(a14) $$\omega_L(\eta) \equiv \left((1-\beta)\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} - (1-\alpha)\frac{s_L(\eta)}{1-s_L(\eta)}\right)\frac{1}{s_L(\eta)}.$$

The expressions for the three derivatives are respectively

(a15) $$\frac{\partial i(\eta)}{\partial \eta} = -\frac{\ln z}{z^\eta}\left(\alpha - \beta\frac{\lambda}{\rho - n + (\sigma-1)g_c + \lambda}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta},$$

(a16) $$\frac{\partial s_K(\eta)}{\partial \eta} = \frac{\beta}{\alpha}\left(\frac{z^\eta \lambda \ln z}{\rho - n + (\sigma-1)g_c + \lambda}\right) \Bigg/ \left(1 + \frac{\beta}{\alpha}\left(\frac{\lambda(z^\eta - 1)}{\rho - n + (\sigma-1)g_c + \lambda}\right)\right)^2,$$

(a17) $$\frac{\partial s_L(\eta)}{\partial \eta} = \frac{1-\beta}{1-\alpha}\left(\frac{z^\eta \lambda \ln z}{\rho - n + (\sigma-1)g_c + \lambda}\right) \Bigg/ \left(1 + \frac{1-\beta}{1-\alpha}\left(\frac{\lambda(z^\eta - 1)}{\rho - n + (\sigma-1)g_c + \lambda}\right)\right)^2. \blacksquare$$

Lemma 4: *In the case of the complete backloading profit-sharing arrangement, the backloading discount factor is given by* $\nu(\eta_{lead}) = (\lambda/(r+\lambda-g_Y))^{\eta_{lead}}$.

Proof: The expected present value of a successful innovation in the case of *complete backloading* is

$$V_t^b = \left(\frac{z^\eta - 1}{z^\eta}\right)\frac{Y_t}{r+\lambda-g_Y} E[e^{-(r-g_Y)s}],$$

where $\eta = \eta_{lead} + \eta_{lag}$. $\eta_{lead} \in \{0, \infty\}$ represents leading breadth, and $\eta_{lag} \in (0,1]$ represents lagging breadth. s is a random variable representing the time it takes for η_{lead} innovations to occur (i.e. when the most recent innovator starts receiving monopolistic profits). Given the Poisson arrival rate of λ, the expected value of s is $E[s] = \eta_{lead}/\lambda$. However, we are interested in $E[e^{-(r-g_Y)s}]$, which is different from $e^{-(r-g_Y)E[s]}$ because of Jensen's inequality. The density function $f(s)$ is needed in order to calculate the expected value. From the Poisson distribution of innovation arrivals, $e^{-\lambda s}(\lambda s)^{\eta_{lead}}/\eta_{lead}!$ is the probability that there are η_{lead} innovations within the time interval s. Therefore, the density function of s, which follows the Erlang distribution, is $f(s) = \lambda^{\eta_{lead}} s^{\eta_{lead}-1} e^{-\lambda s}/(\eta_{lead}-1)!$. Then, the expected value of the *complete-backloading* discount factor is

$$E[e^{-(r-g_Y)s}] = \int_0^\infty e^{-(r-g_Y)s} f(s) ds = \frac{\lambda^{\eta_{lead}}}{(\eta_{lead}-1)!}\left(\int_0^\infty e^{-(r-g_Y+\lambda)s} s^{\eta_{lead}-1} ds\right).$$

Manipulating this expression yields

$$E[e^{-(r-g_Y)s}] = \frac{\lambda^{\eta_{lead}}}{(\eta_{lead}-1)!}\frac{(\eta_{lead}-2)!}{(r-g_Y+\lambda)^{\eta_{lead}-1}}\left(\frac{(r-g_Y+\lambda)^{\eta_{lead}-1}}{(\eta_{lead}-2)!}\int_0^\infty e^{-(r-g_Y+\lambda)s} s^{\eta_{lead}-1} ds\right).$$

The term inside the bracket is the expected value of s from a Gamma distribution with a shape parameter of $\eta_{lead}-1$ and a scale parameter of $1/(r-g_Y+\lambda)$. Therefore, $E[e^{-(r-g_Y)s}]$ simplifies to

$$E[e^{-(r-g_Y)s}] = \frac{\lambda^{\eta_{lead}}}{(\eta_{lead}-1)!}\frac{(\eta_{lead}-2)!}{(r-g_Y+\lambda)^{\eta_{lead}-1}}\frac{\eta_{lead}-1}{r-g_Y+\lambda} = \left(\frac{\lambda}{r-g_Y+\lambda}\right)^{\eta_{lead}}. \blacksquare$$

Appendix II

To introduce the static distortionary effect of markup-pricing into the model, there must be both monopolistic and competitive sectors. To maintain the analytical tractability of the aggregate equations, the CES aggregator is converted to the usual Cobb-Douglas aggregator given by

(b1) $$Y_t = \exp\left(\int_0^1 \ln X_t(j)dj\right).$$

Among the continuum of intermediate goods $j \in [0,1]$, a fraction θ of industries is characterized by perfect competition because innovations in these industries are non-patentable. Without loss of generality, the industries are ordered such that industries $j' \in [0,\theta]$ are competitive. Therefore,

(b2) $$P_t(j') = MC_t(j')$$

for $j' \in [0,\theta]$. The aggregate price level is

(b3) $$P_t = \mu MC_t,$$

where $\mu \equiv (z^{\eta})^{1-\theta}$ is the average markup in the economy, and z^{η} is the markup in the monopolistic industries. The first-order condition from the Cobb-Douglas aggregator implies that the ratio of factor inputs in a competitive industry j' and a monopolistic industry j is

(b4) $$\frac{X_t(j')}{X_t(j)} = \frac{L_{x,t}(j')}{L_{x,t}(j)} = \frac{K_{x,t}(j')}{K_{x,t}(j)} = z^{\eta}.$$

Substituting $X_t(j)$ for $j \in [0,1]$ into (b1), the aggregate production becomes

(b5) $$Y_t = A_t Z_t \left(\frac{K_{x,t}}{L_{x,t}}\right)^{\alpha} L^e_{x,t},$$

where $L^e_{x,t}$ is defined as

(b6) $$L^e_{x,t} \equiv \exp\left(\int_0^1 \ln L_{x,t}(j)dj\right) \neq \left(\int_0^1 L_{x,t}(j)dj\right) = L_{x,t}.$$

In particular, denote $\vartheta(\eta)$ as the ratio of $L^e_{x,t}$ and $L_{x,t}$, which is given by

(b7) $$\vartheta(\eta) \equiv \frac{L^e_{x,t}}{L_{x,t}} = \frac{(z^\eta)^\theta}{z^\eta\theta + (1-\theta)} \in (0,1)$$

for $\theta \in (0,1)$. $\vartheta(\eta)$ represents the static distortionary effect of markup pricing, and it enters the aggregate production function as

(b8) $$Y_t = \vartheta(\eta) A_t Z_t K^\alpha_{x,t} L^{1-\alpha}_{x,t}.$$

The total amount of monopolistic profit is

(b9) $$\pi_t = \left(\int_\theta^1 \pi_t(j) dj \right) = (1-\theta)\left(\frac{z^\eta - 1}{z^\eta} \right) Y_t,$$

and the amount of factor payment to capital in the intermediate-goods sector is

(b10) $$R_t K_{x,t} = \alpha \left(\frac{z^\eta\theta + (1-\theta)}{z^\eta} \right) Y_t.$$

Using the no-arbitrage condition $r_t = R_t - \delta$ and the law of motion for capital, the steady-state rate of investment is

(b11) $$i = \frac{\alpha(z^\eta\theta + (1-\theta))}{z^\eta(1-s_K)} \left(\frac{g_K + \delta}{r + \delta} \right).$$

The aggregate value of patents with a *constant* backloading discount factor ν is

(b12) $$V_t = (1-\theta)\nu \left(\frac{z^\eta - 1}{z^\eta} \right) \frac{Y_t}{r + \lambda - g_Y}.$$

The steady-state R&D shares of labor and capital are respectively

(b13) $$\frac{s_L}{1-s_L} = \frac{1-\beta}{1-\alpha} \left(\frac{\lambda}{r + \lambda - g_Y} \right) \frac{\nu(z^\eta - 1)(1-\theta)}{z^\eta\theta + (1-\theta)},$$

(b14) $$\frac{s_K}{1-s_K} = \frac{\beta}{\alpha} \left(\frac{\lambda}{r + \lambda - g_Y} \right) \frac{\nu(z^\eta - 1)(1-\theta)}{z^\eta\theta + (1-\theta)}.$$

The second-best optimal level of η can be found by maximizing (b15) $$U=\int_0^\infty e^{-(\rho-n)t}\frac{c_t^{1-\sigma}(\eta)}{1-\sigma}dt$$

subject to the aggregate production function given by

(b16) $$Y_t=\vartheta A_t Z_t(1-s_K)^\alpha(1-s_L)^{1-\alpha}K_t^\alpha L_t^{1-\alpha},$$

the law of motion for capital given by

(b17) $$\dot{K}_t=iY_t-K_t\delta,$$

and the law of motion for R&D-driven technology given by

(b18) $$\dot{A}_t=A_t^\phi(s_K)^{\beta\gamma}(s_L)^{(1-\beta)\gamma}K_t^{\beta\gamma}L_t^{(1-\beta)\gamma}\varphi\ln z\,.$$

The current-value Hamiltonian $\tilde{H}$ is

(b19) $$\begin{aligned}\tilde{H}&=(1-\sigma)^{-1}\left(\frac{(1-i)\vartheta A_t Z_t(1-s_K)^\alpha(1-s_L)^{1-\alpha}K_t^\alpha L_t^{1-\alpha}}{L_t}\right)^{1-\sigma}\\&+\upsilon_K(i\vartheta A_t Z_t(1-s_K)^\alpha(1-s_L)^{1-\alpha}K_t^\alpha L_t^{1-\alpha}-K_t\delta)\\&+\upsilon_A A_t^\phi(s_K)^{\beta\gamma}(s_L)^{(1-\beta)\gamma}K_t^{\beta\gamma}L_t^{(1-\beta)\gamma}\varphi\ln z\end{aligned}\quad.$$

Note that $\vartheta(\eta)$, $s(\eta)$, $s_K(\eta)$ and $s_L(\eta)$ are all functions of η. The first-order conditions are

(b20) $$\tilde{H}_K=\frac{\alpha}{K_t}\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma}+\upsilon_K\left(\alpha\frac{i(\eta)Y_t}{K_t}-\delta\right)+\upsilon_A\left(\beta\gamma\frac{\dot{A}_t}{K_t}\right)=(\rho-n)\upsilon_K-\dot{\upsilon}_K,$$

(b21) $$\tilde{H}_A=\frac{1}{A_t}\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma}+\upsilon_K\left(\frac{i(\eta)Y_t}{A_t}\right)+\upsilon_A\left(\phi\frac{\dot{A}_t}{A_t}\right)=(\rho-n)\upsilon_A-\dot{\upsilon}_A,$$

(b22) $$\begin{aligned}\tilde{H}_\eta&=-\left(\frac{(1-i(\eta))Y_t}{L_t}\right)^{1-\sigma}\left(\frac{1}{1-i(\eta)}\frac{\partial i(\eta)}{\partial\eta}-\frac{1}{\vartheta(\eta)}\frac{\partial\vartheta(\eta)}{\partial\eta}+\frac{\alpha}{1-s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta}+\frac{1-\alpha}{1-s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}\right)\\&+\upsilon_K Y_t\left(\frac{\partial i(\eta)}{\partial\eta}+\frac{1}{\vartheta(\eta)}\frac{\partial\vartheta(\eta)}{\partial\eta}i(\eta)-\frac{\alpha}{1-s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta}i(\eta)-\frac{1-\alpha}{1-s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}i(\eta)\right)\\&+\upsilon_A\dot{A}_t\left(\frac{\beta\gamma}{s_K(\eta)}\frac{\partial s_K(\eta)}{\partial\eta}+\frac{(1-\beta)\gamma}{s_L(\eta)}\frac{\partial s_L(\eta)}{\partial\eta}\right)=0\end{aligned}\quad.$$

(b20) and (b21) simplify to the following conditions (b23) $$\upsilon_K K_t(\rho + g_c\sigma + \delta - \alpha(g_K + \delta)) = \alpha((1-i(\eta))Y_t / L_t)^{1-\sigma} + \upsilon_A A_t g_A \beta\gamma ,$$

(b24) $$\upsilon_A A_t(\rho - n + (\sigma-1)g_c + (1-\phi)g_A) = ((1-i(\eta))Y_t / L_t)^{1-\sigma} + \upsilon_K K_t(g_K + \delta) .$$

Using (b23) and (b24), the first-order condition (b22) that characterizes the second-best optimal level of patent breadth simplifies to

(b25) $$\omega_\theta(\eta)\frac{\partial \vartheta(\eta)}{\partial \eta} + \omega_I(\eta)\frac{\partial i(\eta)}{\partial \eta} + \omega_K(\eta)\frac{\partial s_K(\eta)}{\partial \eta} + \omega_L(\eta)\frac{\partial s_L(\eta)}{\partial \eta} = 0 .$$

The ω's are defined as

(b26) $$\omega_\theta(\eta) \equiv \frac{1}{\vartheta(\eta)} ,$$

(b27) $$\omega_I(\eta) \equiv \begin{pmatrix} \left(\alpha + \beta\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta} \\ -\frac{i(\eta)}{1-i(\eta)}\left(1 - \left(\alpha + \beta\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta}\right) \end{pmatrix}\frac{1}{i(\eta)} ,$$

(b28) $$\omega_K(\eta) \equiv \left(\beta\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} - \alpha\frac{s_K(\eta)}{1-s_K(\eta)}\right)\frac{1}{s_K(\eta)} ,$$

(b29) $$\omega_L(\eta) \equiv \left((1-\beta)\frac{\gamma g_A}{\rho - n + (\sigma-1)g_c + (1-\phi)g_A} - (1-\alpha)\frac{s_L(\eta)}{1-s_L(\eta)}\right)\frac{1}{s_L(\eta)} .$$

The expressions for the four derivatives are respectively

(b30) $$\frac{\partial \vartheta(\eta)}{\partial \eta} = -\theta \ln z\left(\frac{z^\eta}{z^\eta\theta + (1-\theta)} - 1\right)\frac{(z^\eta)^\theta}{z^\eta\theta + (1-\theta)} ,$$

(b31) $$\frac{\partial i(\eta)}{\partial \eta} = -\frac{(1-\theta)\ln z}{z^\eta}\left(\alpha - \beta\frac{\lambda\nu}{\rho - n + (\sigma-1)g_c + \lambda}\right)\frac{g_K + \delta}{\rho + g_c\sigma + \delta} ,$$

(b32) $$\frac{\partial s_K(\eta)}{\partial \eta} = \frac{\beta}{\alpha}\left(\frac{\nu(1-\theta)z^\eta\lambda\ln z}{\rho - n + (\sigma-1)g_c + \lambda}\right)\Bigg/\left(z^\eta\theta + (1-\theta) + \left(\frac{\beta}{\alpha}\right)\frac{\lambda\nu(z^\eta - 1)(1-\theta)}{\rho - n + (\sigma-1)g_c + \lambda}\right)^2 ,$$

(b33) $$\frac{\partial s_L(\eta)}{\partial \eta}=\frac{1-\beta}{1-\alpha}\left(\frac{\nu(1-\theta)z^{\eta}\lambda\ln z}{\rho-n+(\sigma-1)g_c+\lambda}\right)\Bigg/\left(z^{\eta}\theta+(1-\theta)+\left(\frac{1-\beta}{1-\alpha}\right)\frac{\lambda\nu(z^{\eta}-1)(1-\theta)}{\rho-n+(\sigma-1)g_c+\lambda}\right)^2.$$

The balanced-growth path of per capita consumption (in log) can be written as

(b34) $$\ln c_t=\ln c_0+g_c t.$$

$g_c t$ represents the balanced-growth path of consumption and is exogenous because of the semi-endogenous growth formulation. The balanced-growth level of per capital consumption at time 0 is

(b35) $$c_0=\tilde{\vartheta}(1-i)(1-s_K)^{\alpha}(1-s_L)^{1-\alpha}A_0Z_0\left(\frac{K_0}{L_0}\right)^{\alpha},$$

where Z_0 is normalized to one. The capital-labor ratio K_0/L_0 and the level of R&D-driven technology A_0 at time 0 are respectively

(b36) $$\left(\frac{K_0}{L_0}\right)^{\alpha}=\left(\frac{\tilde{\vartheta}i\,(1-s_K)^{\alpha}(1-s_L)^{1-\alpha}A_0}{g_K+\delta}\right)^{\alpha/(1-\alpha)},$$

(b37) $$A_0=\left(s_K^{\beta}s_L^{1-\beta}\left(\frac{K_0}{L_0}\right)^{\beta}\right)^{\gamma/(1-\phi)}\left(\frac{\varphi\ln z}{g_A}\right)^{1/(1-\phi)}.$$

After dropping the exogenous growth path and some constant terms, the expression for the balanced-growth level of per capita consumption that depends on $\tilde{\vartheta}(\eta)$, $i(\eta)$, $s_K(\eta)$ and $s_L(\eta)$ is

(b38) $$c_0(\eta)=\left(\begin{array}{l}\tilde{\vartheta}(\eta)^{1/(1-\alpha)}\,i(\eta)^{\frac{\alpha(1-\phi)+\beta\gamma}{(1-\alpha)(1-\phi)-\beta\gamma}}(1-i(\eta))\\ s_K(\eta)^{\frac{\beta\gamma}{(1-\alpha)(1-\phi)-\beta\gamma}}(1-s_K(\eta))^{\frac{\alpha(1-\phi)}{(1-\alpha)(1-\phi)-\beta\gamma}}\\ s_L(\eta)^{\frac{(1-\beta)\gamma}{(1-\alpha)(1-\phi)-\beta\gamma}}(1-s_L(\eta))^{\frac{(1-\alpha)(1-\phi)}{(1-\alpha)(1-\phi)-\beta\gamma}}\end{array}\right).$$

Appendix III

This appendix provides the details of transforming the variables in the differential equations (65) – (68) into their stationary forms for the purpose of computing the transition dynamics numerically. The Euler equation is given by

(c1) $$\dot{c}_t = c_t(r_t - \rho)/\sigma .$$

Define a stationary variable $\tilde{c}_t \equiv c_t/(A_t Z_t)^{1/(1-\alpha)}$, and its resulting law of motion is

(c2) $$\frac{\dot{\tilde{c}}_t}{\tilde{c}_t} = \frac{1}{\sigma}(r_t - \rho) - \frac{1}{1-\alpha}(\lambda_t \ln z + g_Z) .$$

The law of motion for capital accumulation is given by

(c3) $$\dot{K}_t = Y_t - C_t - K_t\delta .$$

Define a stationary variable $k_t \equiv K_t/(L_t(A_t Z_t)^{1/(1-\alpha)})$, and its resulting law of motion is

(c4) $$\frac{\dot{k}_t}{k_t} = (1 - s_{r,t})k_t^{\alpha-1} - \frac{\tilde{c}_t}{k_t} - (\delta + n) - \frac{1}{1-\alpha}(\lambda_t \ln z + g_Z) .$$

The law of motion for the value of patents is given by

(c5) $$\dot{V}_t = (r_t + \lambda_t)V_t - \nu\left(\frac{\mu - 1}{\mu}\right)Y_t .$$

Define a stationary variable $\tilde{v}_t \equiv V_t/(L_t(A_t Z_t)^{1/(1-\alpha)})$, and its resulting law of motion is

(c6) $$\frac{\dot{\tilde{v}}_t}{\tilde{v}_t} = (r_t + \lambda_t) - (1 - s_{r,t})\nu\left(\frac{\mu - 1}{\mu}\right)\frac{k_t^{\alpha}}{\tilde{v}_t} - n - \frac{1}{1-\alpha}(\lambda_t \ln z + g_Z) .$$

The law of motion for R&D-driven technology is given by

(c7) $$\dot{A}_t = A_t\lambda_t \ln z .$$

Define a stationary variable $a_t \equiv k_t^{\alpha\gamma} A_t^{\alpha\gamma/(1-\alpha)-(1-\phi)} Z_t^{\alpha\gamma/(1-\alpha)} L_t^{\gamma}\varphi$, and its resulting law of motion is

(c8) $$\frac{\dot{a}_t}{a_t} = \alpha\gamma(1 - s_{r,t})k_t^{\alpha-1} - \alpha\gamma\frac{\tilde{c}_t}{k_t} - (1-\phi)\lambda_t \ln z + (n\gamma - \alpha\gamma(\delta + n)) .$$

To close this system of differential equations, the endogenous variables (r_t, $s_{r,t}$, λ_t) are also expressed in terms of the four newly defined stationary variables. The interest rate is

(c9) $$r_t = \alpha k_t^{\alpha-1} / \mu - \delta .$$

From the first-order condition of the R&D sector, the share of factor inputs in R&D is

(c10) $$s_{r,t} = (a_t \tilde{v}_t \mu)^{1/(1-\gamma)} / k_t^{\alpha/(1-\gamma)} .$$

From the law of motion of R&D-driven technology, the Poisson arrival rate of innovations is

(c11) $$\lambda_t = s_{r,t}^{\gamma} a_t .$$

Finally, the steady-state values of the variables are

(c12) $$\lambda = g_A / \ln z ,$$

(c13) $$\frac{s_r}{1-s_r} = \frac{v(\mu-1)\lambda}{\rho - n + (\sigma-1)g_c + \lambda},$$

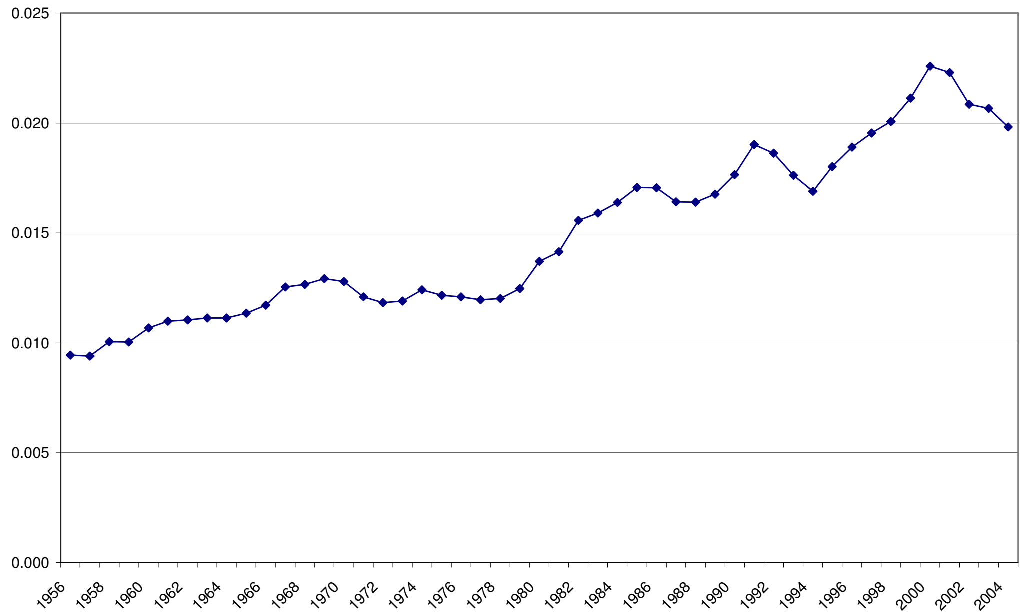
(c14) $$a = \lambda / s_r^{\gamma} .$$

(c15) $$k = \left(\frac{\alpha}{\mu(\delta + \rho + \sigma(g_A + g_Z)/(1-\alpha))} \right)^{1/(1-\alpha)},$$

(c16) $$\tilde{c} = (1-s_r)k^{\alpha} - k\left(\delta + n + \frac{g_A + g_Z}{1-\alpha} \right),$$

(c17) $$\tilde{v} = \frac{v k^{\alpha}(1-s_r)(\mu-1)}{\alpha k_t^{\alpha-1} + \mu\left(\frac{1}{\ln z} - \frac{1}{1-\alpha} \right) g_A - \mu\left(\frac{g_Z}{1-\alpha} + \delta + n \right)} .$$

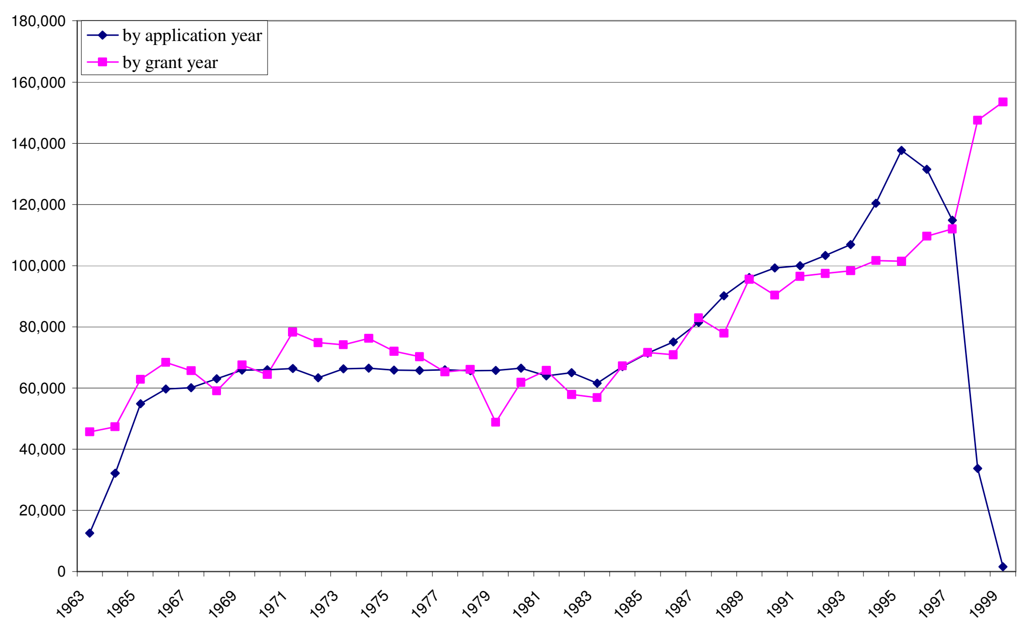
Figure 1: Private Spending on R&D as a Share of GDP



Data Sources: (a) Bureau of Economic Analysis: National Income and Product Accounts Tables; and (b) National Science Foundation: Division of Science Resources Statistics.

Footnote: R&D is net of federal spending, and GDP is net of government spending.

Figure 2: Number of Patents Granted



Data Source: Hall, Jaffe and Trajtenberg (2002): The NBER Patent Citation Data File.

Figure 3: First-Best Optimal R&D Shares for Different Values of ξ and γ

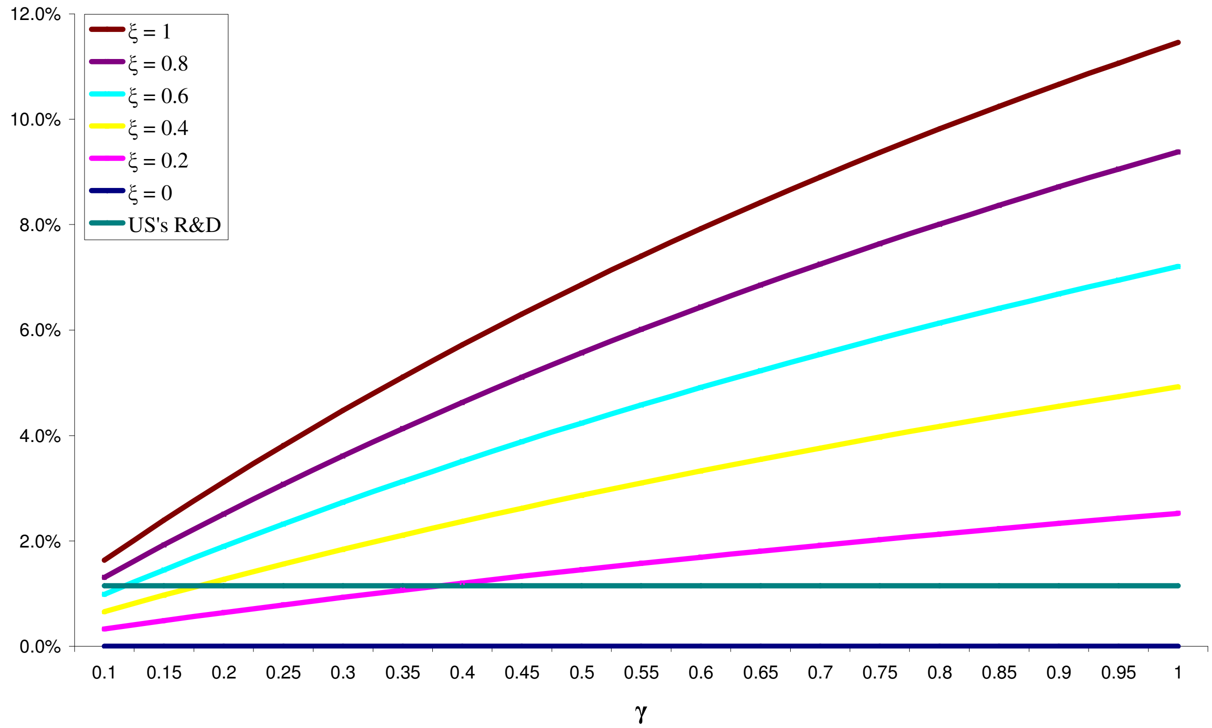


Figure 4: Percentage Change in Long-Run Consumption from Eliminating Blocking Patent

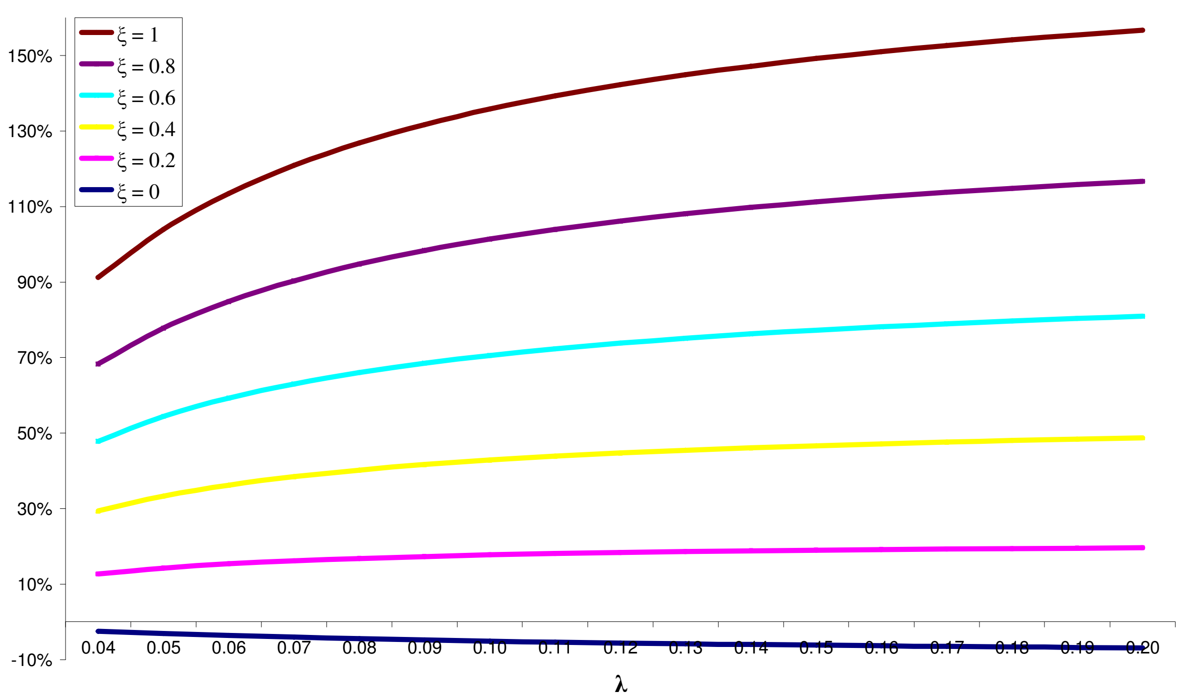


Figure 5a: Optimal Markup for ε = 0.8

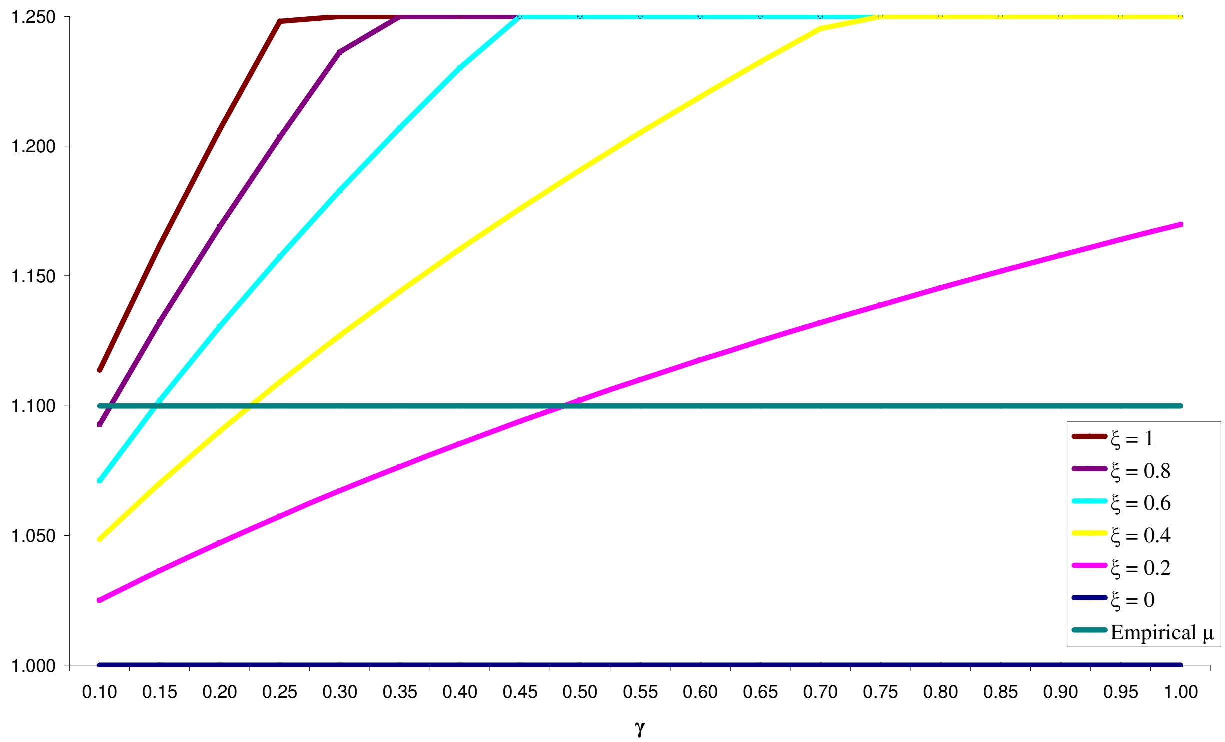


Figure 5b: Optimal Markup for ε = 0

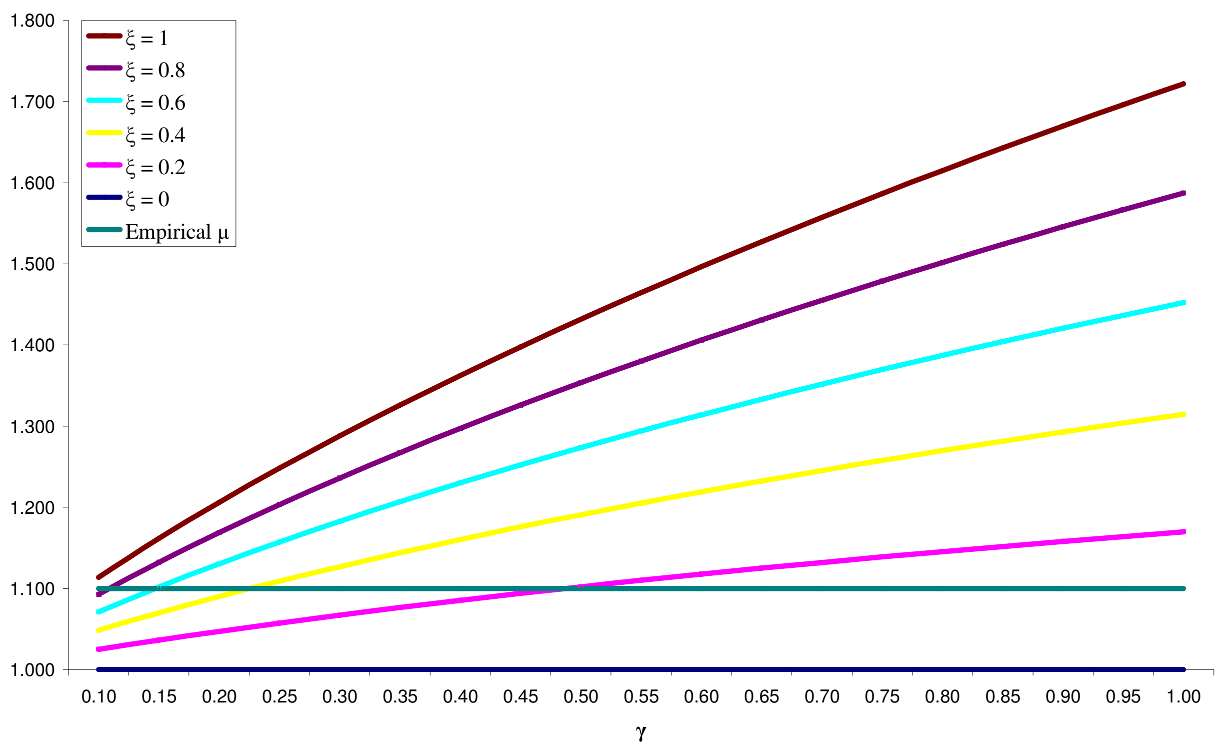


Figure 6a: Percentage Change in Long-Run Consumption for ε = 0.8

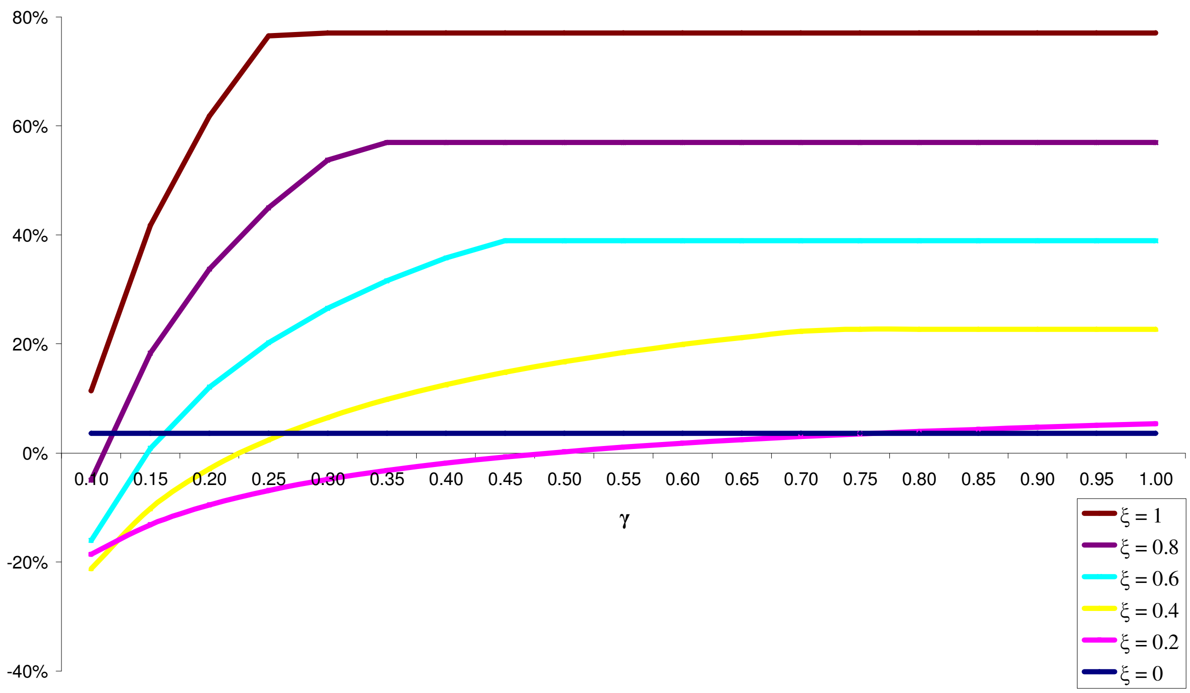


Figure 6b: Percentage Change in Long-Run Consumption for ε = 0

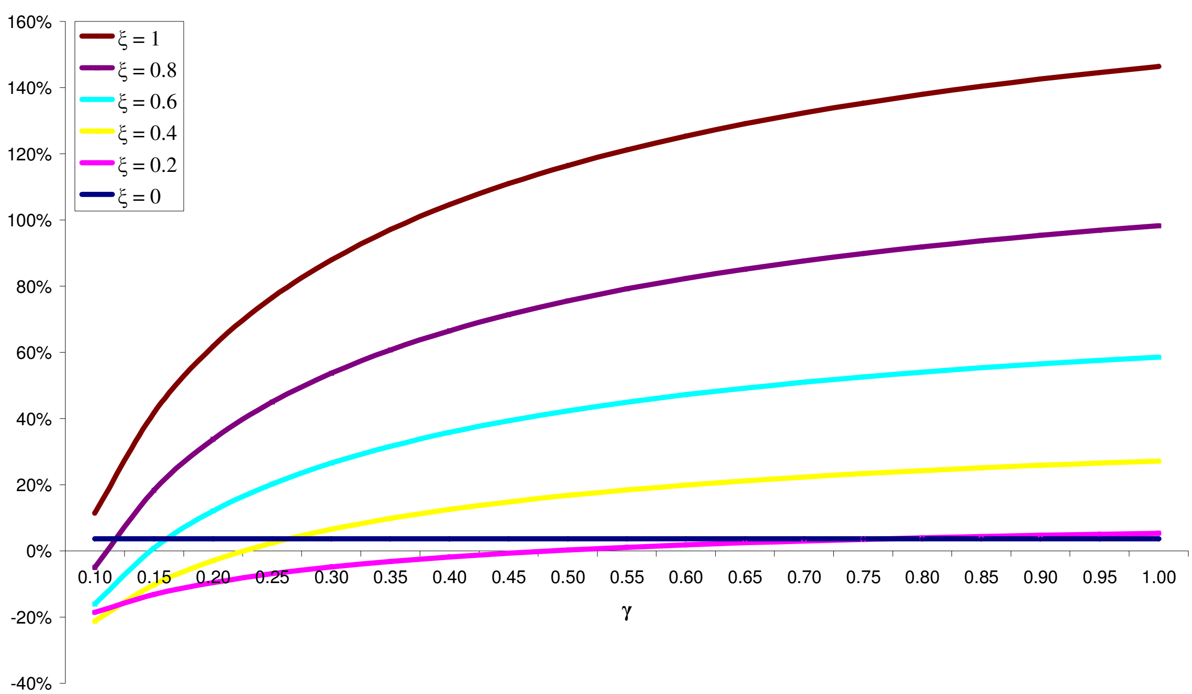


Figure 7a: Optimal Markup for θ = 0.25

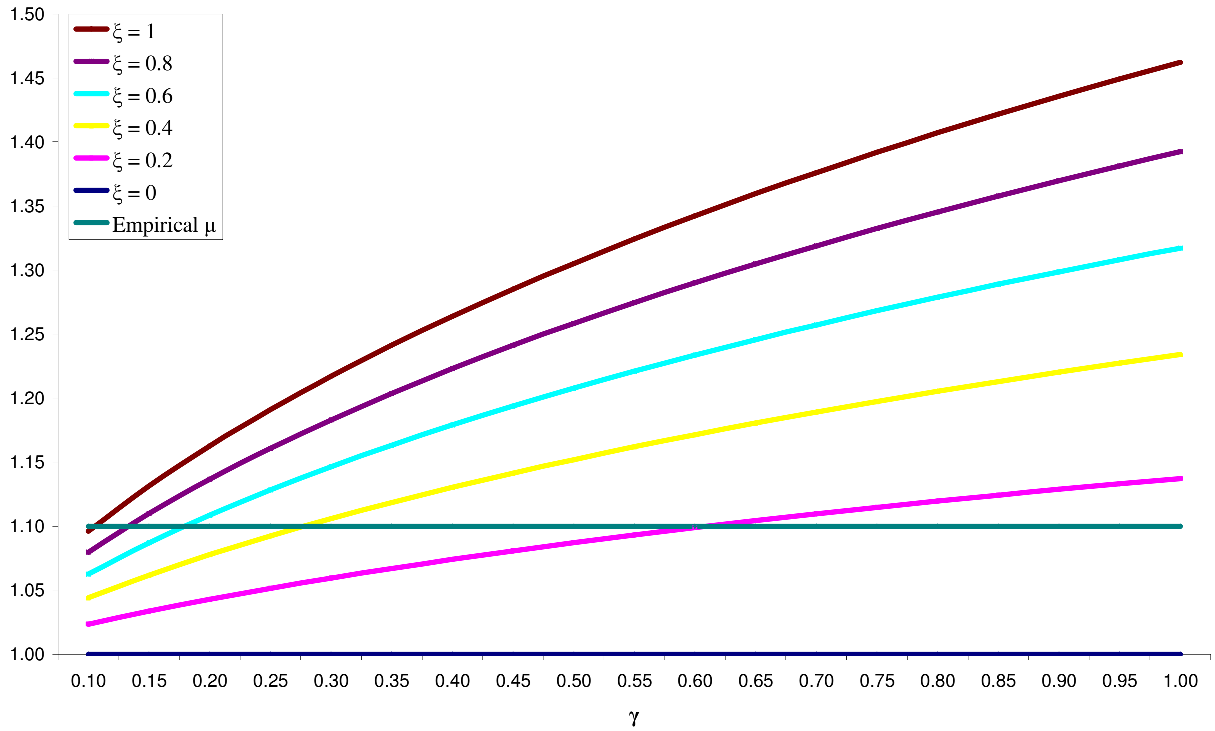


Figure 7b: Optimal Markup for θ = 0.5

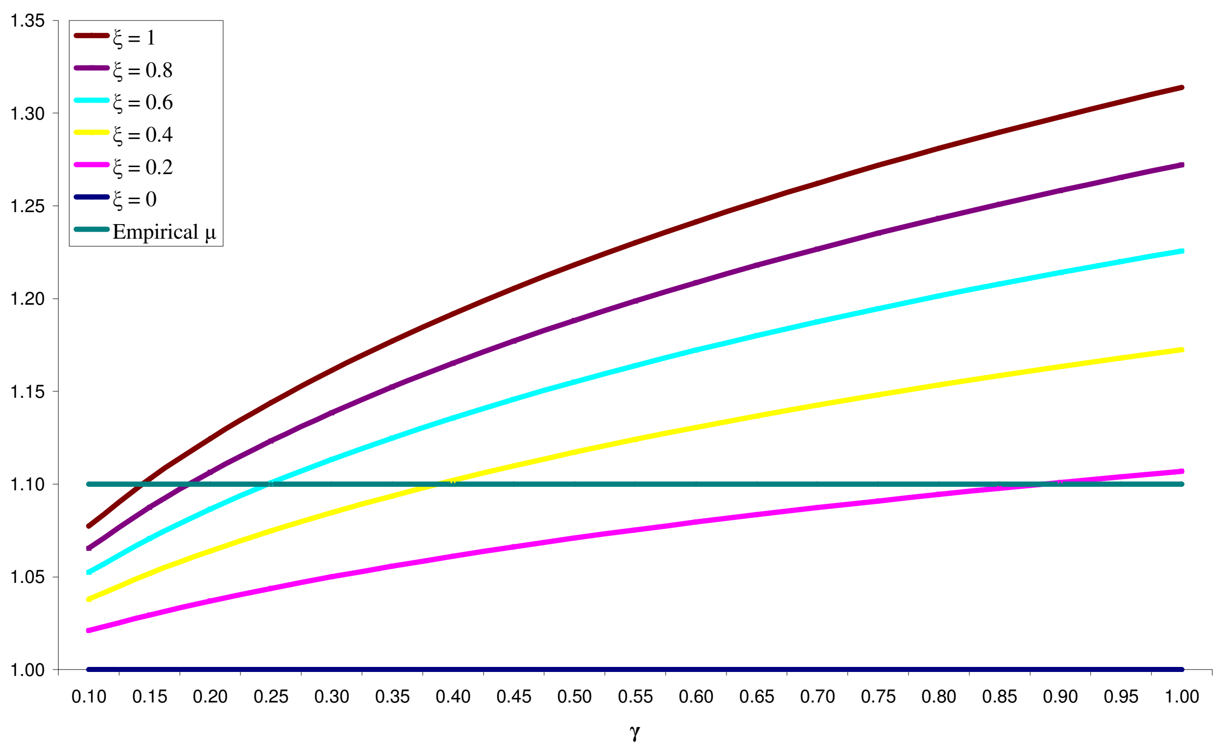


Figure 8a: Percentage Change in Long-Run Consumption for θ = 0.25

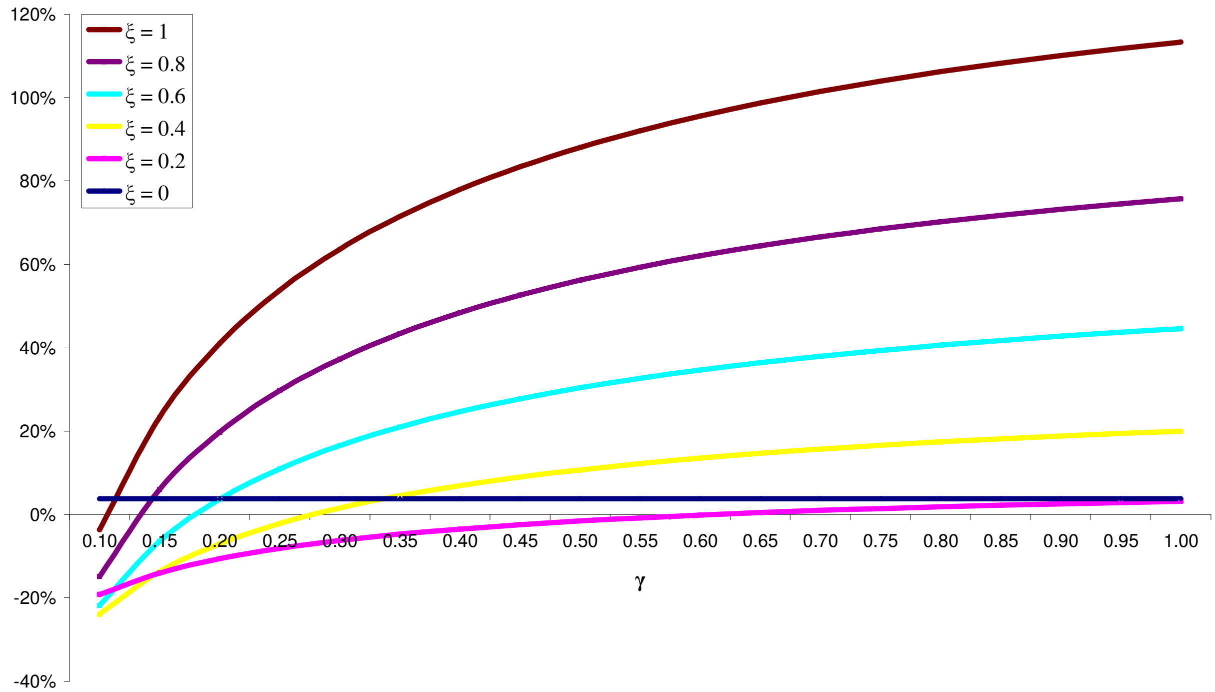


Figure 8b: Percentage Change in Long-Run Conumption for θ = 0.5

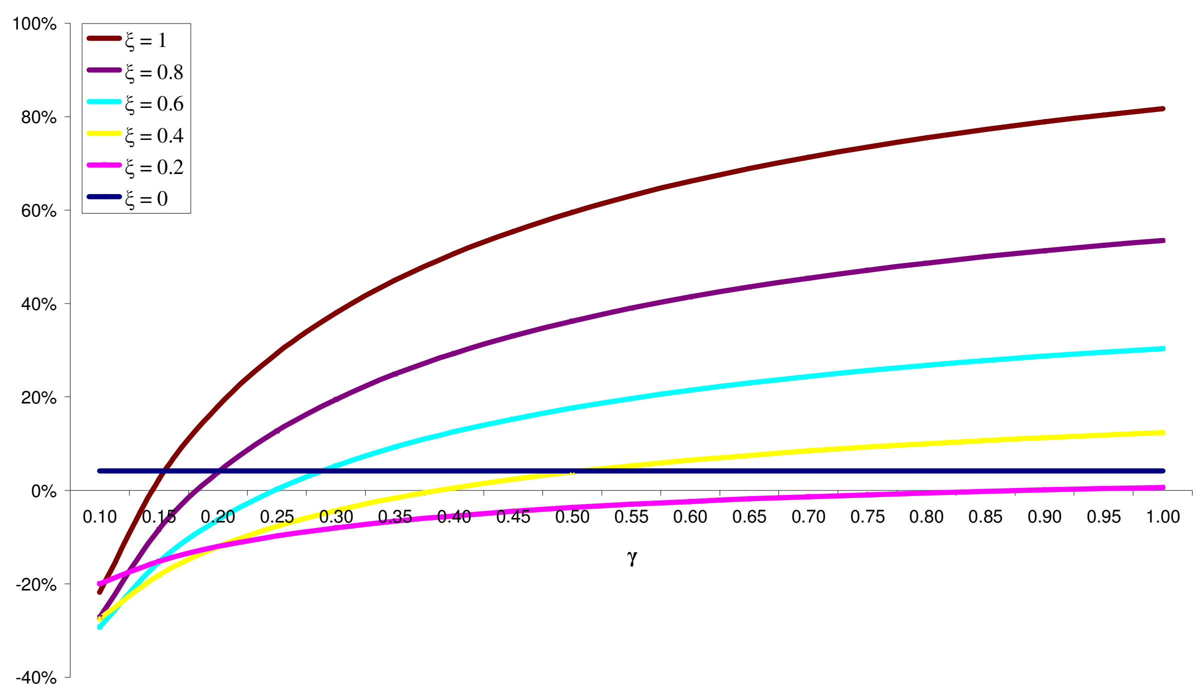


Figure 9a: Transition Dynamics of Log Consumption with δ = 0.08

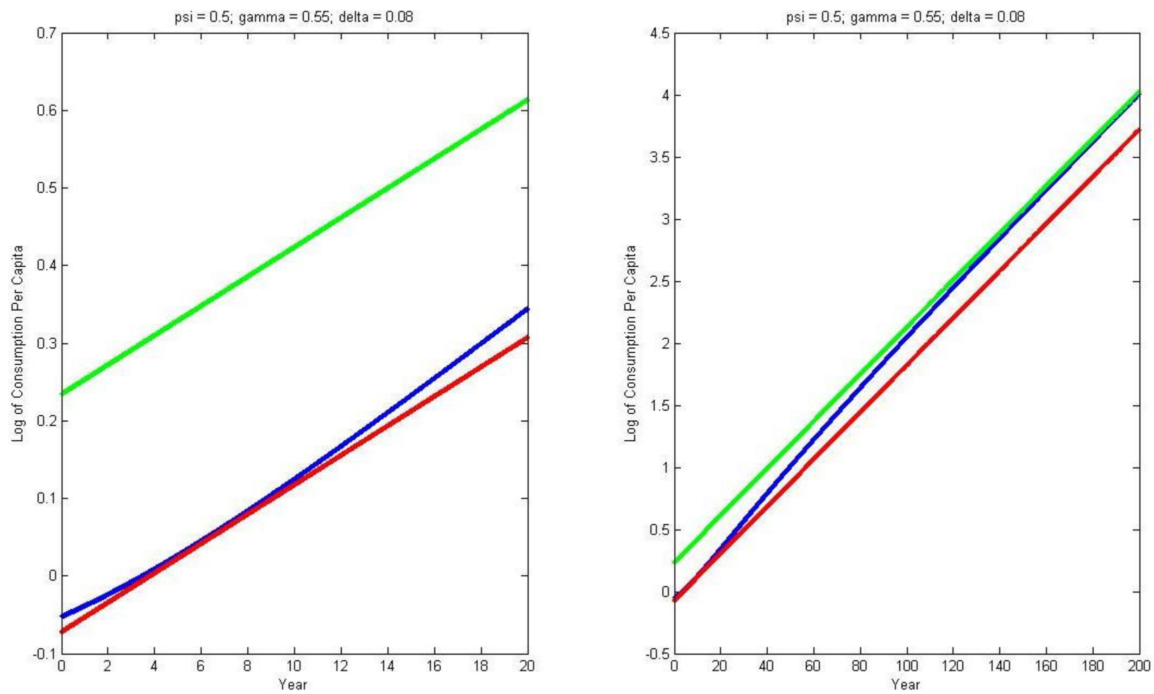


Figure 9b: Transition Dynamics of Log Consumption with δ = 1