



Munich Personal RePEc Archive

Optimal Patent Life in a Variety-Expansion Growth Model

Lin, Hwan C.

University of North Carolina at Charlotte

6 August 2013

Online at <https://mpra.ub.uni-muenchen.de/49790/>

MPRA Paper No. 49790, posted 13 Sep 2013 08:02 UTC

Optimal Patent Life in a Variety-Expansion Growth Model

Hwan C. Lin*

August 2013

Abstract

This paper presents more channels through which the optimal patent life is determined in a R&D-based endogenous growth model that permits growth of new varieties of consumer goods over time. Its modeling features include an endogenous hazard rate facing incumbent monopolists, the prevalence of research congestion, and the aggregate welfare importance of product differentiation. As a result, a patent's effective life is endogenized and less than its legal life. The model is calibrated to a global economy with a set of baseline parameter values. Under the benchmark patent length of 20 years, the calibrated model can deliver along the balanced growth path a plausible innovation rate of 2.84% per year and an economy-wide markup rate of 1.15. The optimal patent length is computed with the algorithm of Golden Search Section, ranging from 17 to 19 years. With the creative-destruction hazard, the world needs a *longer* patent term to maximize social welfare, but with the prevalence of research congestion, the world needs a *shorter* patent term. However, if the world's aggregate welfare appreciates varieties of goods in a way strong enough, the optimal patent term can surprisingly extend beyond even 1,000 years!

Keywords: patent length, innovation, creative destruction, endogenous hazard rate

JEL Classification: O31, O34

© All rights reserved.

*Department of Economics, Belk College of Business, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, U.S.A. Email: hwlin@uncc.edu

1 Introduction

Innovation of new products or production processes is critical to a country's long-term economic growth and standard of living. But innovating activities are risky and costly. They would not be viable if there were no systematic mechanism to prevent free-riders from infringing on intellectual property rights. To fix the free-rider problem, the patent system has been in place for a long time in the world.¹ It grants the inventor (the patentee) for a limited period of time the exclusive right to prevent others from making, using, selling, or distributing the patented invention without permission. The patent term (or "patent length") determines the duration, or the life, of patents for which the inventor is entitled to the exclusive right. Today, under the WTO (World Trade Organization) Agreement on TRIPS (Trade-Related Aspects of Intellectual Property Rights), the length of patents gives a patent life of at least twenty years.²

However, it is well-known that the patent system is merely a second-best policy instrument. Patents provide the inventor with a purposive incentive to innovate, because a successful innovation may generate monopoly profits during the patent life. The world's patent system can therefore generate dynamic social gains from ongoing technological innovation, but at the social cost of static inefficiency arising from monopoly distortions, as patented objects are often sold at monopoly prices.³ Thus, by fine-tuning the term of patents, the socially optimal patent life must be such that balances the dynamic innovation gain against its associated static-inefficiency cost. The present paper formulates more channels through which the optimal patent life is determined for a closed economy.

In a more general setting than the literature on patent length, innovation, and welfare, this paper uses a dynamic general equilibrium model where a patent system permits endogenous growth of new varieties of consumer goods over time.⁴ In such a model, goods are all horizontally differentiated

¹The world's patent system can date back to the fifteen century.

²The WTO Agreement on TRIPS was concluded at the end of the Uruguay Round of the General Agreement on Tariffs and Trade (GATT) in 1994. It sets down minimum standards for many forms of intellectual property rights.

³In general, the patented object may refer to a final product, an intermediate input, or a production process.

⁴Following Nordhaus (1969), earlier studies on the optimal duration of patents were confined to partial equilibrium analysis of one single process innovation, which leads to reduction of marginal production costs. These studies include Scherer (1972), DeBrock (1985), Gilbert and Shapino (1990), and others. In contrast, Judd (1985) represents the initial attempt to study the optimal patent length for a closed economy using a dynamic general equilibrium model where product innovation comes as a stream over time. In Judd's model, the long-term innovation rate is exogenous and the optimal patent length is infinite. In contrast to Judd (1985), Segerstrom et al (1990) develop a quality-ladder endogenous growth model of North-South trade to study the steady-state effects of changing patent length on R&D incentives, technology transfer and relative wages, but how to optimize the patent length is not answered.

and imperfectly substitutable, and innovation of a new good gives birth to a new firm, which is granted by the government a patent with a finite term (say, twenty years). Such a patented firm would continue to be a monopolist, *either* until the patent it holds has expired *or* until it has been forced prematurely to lose monopoly power to newer product innovation. We use the term “creative destruction of a patent’s monopoly profits” to coin the event of a newer product innovation coming to force any patented firms to lose monopoly power before their patents expire. We endogenize an instantaneous hazard rate as an externality of an economy’s aggregate innovating activities, as in Jones and Williams (2000).⁵ Following the well-known survival analysis, We use the hazard rate to derive a survival function that determines the instantaneous probability that an incumbent may get around the creative-destruction hazard. In the modeled economy, if product innovation proceeds at a faster rate, any incumbent monopoly firm faces a greater hazard of being “creatively destructed” at any point in time, even though the firm still holds a legally valid patent.

Thus, with the hazard rate and its associated survival function coming as an externality of aggregate innovation, any patented firm’s expected monopoly duration is endogenous in the modeled economy and must be *shorter* than its patent’s remaining life to maturity at any point in time.⁶ This important feature distinguishes the present paper from previous studies on patent policy using R&D-based variety-expansion growth models. For instance, in a class of variety-expansion models [e.g., Judd (1985), Chou and Shy (1991), Iwasiako and Futagami (2003) and Grossman and Lai (2004), Futagami and Iwasiako (2007), Chu (2010), and Grinols and Lin (2011)], a patented firm’s expected monopoly duration is always tied to its patent’s remaining life to maturity, since these cited studies all presume a zero hazard rate. In another class of such models [e.g., Helpman (1993), Grinols and Lin (1997, 2006), Lin (2002), and Furukawa (2007)], the hazard rate is an exogenous

⁵Other works endogenizing the hazard rate as an externality of aggregate innovating activities are mostly quality-ladder growth models including Grossman and Helpman (1992, ch.4), O’Donoghue and Zweimüller (2004), and Iwaisako, et al. (2011), among others. There are also works that endogenize the hazard rate *as a purposive result* of profit-driven activities, respectively, in imitation, outbound FDI (foreign direct investment), and rent seeking-cum-innovation. These related works will be discussed in subsection 2.1.3.

⁶For illustrative purposes, let the hazard rate of creative destruction stay constant at $h \in (0, 1)$ and let the patent length l be a positive integer on a discrete time axis $t = -\infty, \dots, -2, -1, 0, 1, 2, \dots, \infty$. Then for an innovative firm that was granted a patent at time $\tau \in (t - l, t]$, its expected monopoly-power duration at time t is equal to $\sum_{s=0}^{l-(t-\tau)-1} (1-h)^s = 1 + (1-h) + (1-h)^2 + \dots + (1-h)^{l-(t-\tau)-1} = \frac{1-(1-h)^{l-(t-\tau)}}{1-(1-h)}$, which is always less than patent’s remaining life $h - (t - \tau)$, provided $0 < h < 1$. If the patent length is infinite ($l \rightarrow \infty$), then the expected monopoly-power duration is simply equal to $1/h$. The greater the hazard rate h , the shorter the expected monopoly-power duration. As will be explained in the text, the hazard rate h is positively associated with the rate of product innovation. Note that in the continuous-time, the hazard rate h is not a probability, and so it does not have to be less than one. This will be made clearer in subsection 2.1.3.

fraction but patents are infinitely lived. Therefore, any patented firm's expected monopoly duration is always fixed at the inverse of the hazard rate at any point in time. In contrast to the two classes of variety-expansion models, the present study lays out a modeled economy in which the patent length is finite, the hazard rate is endogenous, and patents are of different vintage. Thus, in this economy, the market value of patents can be determined through more avenues including the cost-of-capital elements (interest rate, hazard rate, innovation rate) and the patent length, as will be explained later.

In addition, we introduce another two modeling features. One is to reflect the empirical prevalence of research congestion and patent races, for which many potential innovators engage in similar research projects and the innovator who successfully races to receive a patent may force the other rivals to abandon in-progress projects, thereby creating wasteful duplication of effort.⁷ The other is to capture the welfare implication of "aggregate importance of product differentiation" by allowing for a wider parameter space that governs the curvature of flow utility with respect to growth of available goods.⁸ Thus, in the present study, the dynamic effects of R&D are more multifaceted: on the one hand, more active R&D leads to faster product innovation and contributes to the stock of knowledge capital, but on the other hand, a faster pace of product innovation must raise the hazard rate (and thereby shorten any patented firm's expected monopoly duration), create wastes of duplicated research, and even deepen aggregate R&D intensity.

Another noteworthy aspect of the present paper is that: an ongoing endogenous innovation process influences the hazard rate, which in turn determines a patent's effective patent life. In a sense, the present paper actually deals with endogenous intellectual property rights (IPR) protection. Though a patent's legal life is exogenous, its effective life is endogenized by the innovation-determined hazard rate in a dynamic general equilibrium framework. Therefore, this paper is an important supplement to the endogenous IPR literature that focuses on rent-seeking activities; see,

⁷See Wright (1983) for an earlier criticism of the patent mechanism that may spur wasteful races.

⁸In the real world the same good does not always equally contribute to household's well-being in different regions. To deal with this problem, Grinols and Lin (1997, 2006) use R&D-based growth models to introduce bi-directional innovation of goods targeted to regional consumers in examining issues regarding North-South protection of intellectual property rights. Daiwan and Rodrik (1991) also address the similar problem, but in a static North-South model. In the vast literature on patent policy, however, goods are often assumed to be equally substitutable in consumption or production, regardless of the location of consumers. Thus, social welfare functions used in this literature tend to overstate the aggregate importance of product differentiation, if regional consumers actually demand regionally-differentiated goods. Following Grinols and Lin (2011), the present study uses a utility function that can discount the aggregate importance of product differentiation by adjusting the value of a preference parameter.

for instance, Dinopoulos and Syropoulos (2007), Eicher and García-Peñalosa (2008), Grieben and Sener (2009), Davis and Sener (2012).

We calibrated the model of the paper to a global economy with a set of baseline parameter values. Under the benchmark patent length of twenty years, the calibrated model can deliver along the balanced growth path a plausible innovation rate equal to $g = 2.84\%$ per year, a fraction of monopoly-supplied (patented) goods about equal to $\zeta_p^M = 22\%$, a creative-destruction hazard rate of $h = 2.84\%$ per year, an effective cost of capital close to 10% per year, and an economy-wide markup rate of 1.15 .^{9,10,11} Simulations indicate that a minimum patent length of more than eleven years is required for the world to keep an active R&D sector. If the patent length were set to be infinite (as opposed to twenty years benchmark), the innovation rate would increase (or less-than-double) to $g = 4.92\%$, whereas the fraction of monopoly-supplied goods would increase (or more than double) to $\zeta_p^M = 50\%$ along with a much greater hazard rate of $h = 4.92\%$ and a much greater effective cost of capital of about 14% . Compared to the 20 years benchmark patent length, the infinite patent term makes the world worse off: based on our simulations, the dynamic gain from faster product innovation is actually dominated by the social costs resulting from increased research intensity (which calls for a higher saving rate) and increased monopoly distortions. These results therefore point to a finite patent length that maximizes social welfare in our calibrated model.

The search for such an optimal patent length presents a constrained lifetime welfare maximization problem.¹² We applied an algorithm known as “Golden Section Search” to numerically solve the maximization problem.¹³ We found that the world’s optimal patent length equals 18.35 years under our chosen baseline parameter values and that it may range from 17 to 19 years from

⁹According to OECD’s published annual patent data of 1985 through 2008, the average annual growth rate of the world’s triadic patent families was about 3% , which may reveal the world’s annual innovation rate in general. Triadic patent families refer to a set of patents taken from the European Patent Office (EPO), the Japanese Patent Office (JPO), and U.S. Patent and Trademark Office (USPTO) that share one or more priorities (see Dernis and Khan 2004).

¹⁰In the paper we assumed the annual rate of time preference equal to $\rho = 4\%$. Along a balanced growth path, the cost of capital equals $\rho + h = 6.84\%$. But ongoing product innovation is seen to dilute an incumbent monopoly firm’s profit flow in a more crowded market. With the steady-state innovation rate of 2.84% , the effective cost of capital is equal to $\rho + h + g = 9.68\%$ per year. Detailed derivations are given in the text.

¹¹The economy-wide markup rate is endogenous in the paper. The calibrated model of the paper delivers a markup rate that falls onto the empirical estimates of 1.05 to 1.40 ; see Norrbin (1993) and Basu (1996).

¹²In the modeled economy, the patent length does not directly enter the objective (social welfare) function, but exerts its effects through a steady-state system of two nonlinear equations, which, however, do not have a closed form solution.

¹³The algorithm of golden section search is the counterpart of bisection used in finding roots of equations; see Kiusalaas (2010, pp. 376-379).

our sensitivity analysis. Our robustness checks suggest that with the creative-destruction hazard confronting incumbent firms, the world needs a *longer* patent term to maximize social welfare. But with the research congestion effect, the world instead needs a *shorter* patent term. These interesting findings make intuitive sense: the hazard of creative destruction reduces a patented firm's expected monopoly duration and therefore calls for a longer optimal patent term, while the research congestion effect raises the social cost of research intensity and therefore calls for a shorter optimal patent term. However, our robustness checks also demonstrate that the optimal patent term appears quite sensitive to the aggregate welfare importance of product differentiation: if aggregate welfare appreciates varieties of goods in a way strong enough, the optimal patent term can surprisingly extend beyond even 1,000 years!

The paper contributes to the literature on patent length, social welfare, and an aggregate economy. Its numerical simulations based on a calibrated growth model with expanding variety stands in contrast to some closely related previous work. The seminar paper of Judd (1985) shows that the socially optimal patent length is infinite for a closed economy with an exogenous long-run innovation rate. Chou and Shy (1991) show that a finite patent length may be optimal if the cost of developing a new product is increasing with the flow of product development. Their model, however, does not consider household's intertemporal decisions on saving and investment.¹⁴ Using an endogenous growth model similar to Romer (1990), Iwasiako and Futagami (2003) present an analytic proof that a finite duration of patents can be optimal, while Futagami and Iwasiako (2007) show that an infinite life of patents is not socially optimal in the absence of the scale effects (Jones 1995). Yet, their works do not quantify the optimal patent length, and do not allow to endogenize effective patent life, since they assume patented firms can always survive the entire statutory patent life.

The rest of the paper is organized as follows. Section 2 lays out the model for a closed economy in which the representative forward-looking household supplies saved funds to finance R&D (research & development) and the forward-looking innovators seek to earn their expected intertemporal profits by engaging in R&D. Section 3 calibrates the model to a global economy and conducts computational simulations to see how changing the patent length can affect product innovation and

¹⁴In Chou and Shy (1991), the interest rate is exogenous, and the profit flow of a patented firm and the value of patents are both time-invariant in steady state.

other variables along the balanced growth path. Robustness checks on some structural parameters are also conducted. Section 4 computes the optimal duration of patents for our calibrated model using the golden section search algorithm. Lastly, Section 5 concludes.

2 The Model

In this section we develop a dynamic general-equilibrium model for a closed economy, where the representative household optimizes consumption/saving flows to maximize her lifetime utility under a dynamic budget constraint. The saving flow is then channeled through financial markets into profit-oriented R&D activities that introduce newer varieties of innovated goods over time. In this economy a patent system is in place to protect intellectual property rights to R&D-driven innovations. The government issues patents to protect all innovative goods, so that the patent holder (or called the “innovative firm”) is granted monopoly power over the sale of its patented product, but only for a finite period of time stipulated by a uniform patent length $l \in (0, \infty)$ applicable to all issued patents. Under the patent system, the patent holder will *expectantly* lose monopoly power when the patent expires. But the patent holder may also *prematurely* lose monopoly power to newly emerging innovations, even if the patent has not yet expired. This unexpected event, coined “Schumpeterian creative destruction” in the model, is an endogenous hazard facing all incumbent monopolists, as will be fully explained later. In what follows we lay out the model where saving and investment interact in a forward-looking manner to drive product innovation onwards.

2.1 Product innovation and creative destruction

In the model of this paper $n[t]$ denotes the measure of product availability at time t . Out of the $n[t]$ available products at time t , there are $n_p[t]$ patented products that received patents previously on patent-issuance dates $\tau \in (t-l, t]$ and $n_{np}[t]$ unpatented products that had completely eclipsed their term- l patent life. These unpatented products received patents on patent-issuance dates $\tau \in (-\infty, t-l]$. At any point in time, it holds that $n = n_p + n_{np}$.¹⁵ This equality relationship is equivalent to $\zeta_p + \zeta_{np} = 1$ in fraction terms, where $\zeta_p \equiv \frac{n_p}{n}$ measures the fraction of products holding live patents and $\zeta_{np} \equiv \frac{n_{np}}{n}$ the fraction of products holding dead (expired) patents.

¹⁵Henceforth, unless necessary, time variables may be suppressed for brevity.

2.1.1 Growth rates of patented and unpatented products

Following the convention that $\dot{z} \equiv dz/dt$ throughout the paper, we define $g[t] \equiv \frac{\dot{n}[t]}{n[t]}$ as the instantaneous growth rate of available products (or the innovation rate for brevity), $g_p[t] \equiv \frac{\dot{n}_p[t]}{n_p[t]}$ as the instantaneous growth rate of patented products, and $g_{np}[t] \equiv \frac{\dot{n}_{np}[t]}{n_{np}[t]}$ as the instantaneous growth rate of unpatented products. Then, from $\dot{n} = \dot{n}_p + \dot{n}_{np}$, the innovation rate g must be a weighted average of g_p and g_{np} . That is,

$$g[t] = \zeta_p[t]g_p[t] + \zeta_{np}[t]g_{np}[t]. \quad (1)$$

Differentiating $\zeta_p \equiv \frac{n_p}{n}$ with respect to time t yields

$$\dot{\zeta}_p[t] = (g_p[t] - g[t])\zeta_p[t] = (g[t] - g_{np}[t])\zeta_{np}[t] \quad (2)$$

where $\dot{\zeta}_p[t] = -\dot{\zeta}_{np}[t]$. As indicated by (2), the patented-product fraction ζ_p *increases* while the unpatented-product fraction ζ_{np} *decreases* over time if $g_p > g$ or if $g > g_{np}$. Both ζ_p and ζ_{np} are predetermined stock variables at any point in time. Driven by growth rates g , g_p , and g_{np} , these stocks will change continuously through time. These growth rates, however, are determined by the model's inherent innovation mechanism and institute arrangement (i.e., patent length l), as we will explain below.

2.1.2 Innovation and research productivity

First, the flow of varieties of goods is determined by the R&D production function,

$$\dot{n}[t] = \frac{n[t]L_{R\&D}[t]}{\mu[t]} \quad (3)$$

where $n[t]$ serves as the stock of knowledge, $L_{R\&D}[t]$ is the quantity of labor employed for product innovation, and $\mu[t]$, an inverse measure of research productivity, is endogenously determined by

$$\mu[t] = \frac{\theta}{1 - \phi g[t]}, \quad \theta > 0, \quad \phi > 0 \quad (4)$$

The positive coefficient ϕ captures the prevalence of research congestion, or patent race, or anything else that may decrease research productivity with ongoing expansion of innovative products.¹⁶ If ϕ were set equal to zero, the research productivity measure $\mu[t]$ would be fixed at a constant parameter θ , as in Romer (1990). But given $\phi > 0$, the measure of $\mu[t]$ must vary with the innovation rate $g[t]$. From (3), the knowledge stock accumulation (i.e. increase in $n[t]$) *increases* labor’s R&D productivity in general, but the faster the pace of product innovation (i.e. increase in $g[t]$), the greater the value of $\mu[t]$, and the smaller the research productivity of labor. Thus, while the stock $n[t]$ of knowledge has a positive externality for future innovation, the flow $g[t]$ of knowledge has a negative externality in that it lowers research productivity of labor equipped with the same stock of knowledge. The latter can be interpreted as a wasting of resources via research congestion, patent race, or the like. Substituting (4) into (3) yields

$$g[t] = \frac{L_{R\&D}[t]}{\theta + \phi L_{R\&D}[t]} \quad (5)$$

It becomes clear that g is concave in the amount of R&D labor $L_{R\&D}$, given that $\phi > 0$, so that R&D productivity decreases if more R&D is undertaken. As $\frac{\partial g}{\partial L_{R\&D}} \big|_{L_{R\&D}=0} = \frac{1}{\theta} < \infty$, there is an upper bound to R&D productivity and therefore a lower bound on marginal R&D costs. As a consequence, (5) does not satisfy the Inada condition of $\frac{\partial g}{\partial L_{R\&D}} \big|_{L_{R\&D}=0} \rightarrow \infty$. As we will demonstrate later, this modeling feature calls for a minimal patent length below which no firm finds it profitable to undertake any R&D.¹⁷

Given research employment $L_{R\&D}[t]$, the flow of innovations means that $\dot{n}[t]$ ($= n[t]g[t]$) patents are newly issued at time t , and these patents will expire at time $t + l$: that is, the patent length l ensures $\dot{n}[t] = \dot{n}_{np}[t + l]$, or equivalently, the flow of unpatented products is determined by

¹⁶In precise sense, equation (4) can be written as $\mu[t] = \theta \cdot \max[\frac{1}{1-\phi g[t]}, 0]$ for $\mu \geq 0$. As such, combining (3) and (4) yields $\dot{n}[t] = \frac{n[t]}{\theta} L_{R\&D}[t] \max[\frac{1}{1-\phi g[t]}, 0]$. This production function of newer varieties has been used in Cozzi (2003; see equation(1)). In Cozzi’s interpretation, parameter ϕ is to capture the cost of “monitoring innovation,” thereby permitting Cozzi’s Romer-style model to remove the well-known scale effects, in a way different from Jones (1995). Cozzi and Spinesi (2004) use a slightly modified innovation technology to model both horizontal and vertical innovations, where the stock of knowledge is not represented by the stock of horizontal innovation ($n[t]$) as in equation (3), but rather by a function of the stocks of horizontal and vertical innovations. Regardless, in either this paper, or Cozzi (2003), or Cozzi and Spinesi (2004), the steady-state innovation rate is tied to the inverse of parameter ϕ if population of labor is allowed to grow at a constant rate. This steady-state feature can be seen by turning equation (3) into equation (5). I thank an anonymous referee for pointing out these earlier related papers.

¹⁷This feature particularly matters for pharmaceutical or biological drugs development, which is very much costly in terms of the lengthy clinical trials and their high risks of failures.

$\dot{n}_{np}[t] = \dot{n}[t - l]$. As such, the growth rates of out-of-patent and patented products can be easily derived as¹⁸

$$g_{np}[t] = \left(\frac{1}{\zeta_{np}[t]} \right) g[t - l] e^{-\bar{g}[t]l} \quad (6a)$$

$$g_p[t] = \left(\frac{1}{\zeta_p[t]} \right) (g[t] - g[t - l] e^{-\bar{g}[t]l}) \quad (6b)$$

where $\bar{g}[t] \equiv \frac{1}{l} \int_{t-l}^{t^-} g[s] ds$ represents the average of past innovation rates at $s \in [t - l, t)$. Evidently, $g_{np}[t]$ is predetermined by history due to (6a), but $g_p[t]$ is not in terms of (6b) because $g[t]$ can change at any moment and therefore affects $g_p[t]$ at the same time.

2.1.3 Innovation and the hazard-rate function

The preceding analysis depicts the dynamics of patented and out-of-patent products under a patent system with a finite term of patent life. Under such a system, the government determines the patent length l , which represents a patented firm's maximum potential duration of monopoly power. Any patented firm will expectantly lose monopoly power and relegate to a perfectly competitive firm, as soon as her patent expires. However, in contrast to the literature [Judd (1985), Chou and Shy (1991), Iwasiako and Futagami (2003) and Grossman and Lai (2004), Futagami and Iwasiako (2007), Chu (2010), and Grinols and Lin (2011)], the paper allows for a non-zero hazard rate at which any patent holder may *prematurely* lose monopoly power to newly patented innovators, thereby becoming a competitive producer. We call such an innovation-induced incidence "creative destruction of a patent's monopoly profits." Here the incidence of "creative destruction" shifts profits from an old monopolist to a newly innovative monopolist, while the old one remains on the market by being a competitive producer. Hence, the connotation of "creative destruction" is not exactly the same as in most R&D-based endogenous growth models, where an incidence of creative destruction not only shifts profits to new innovators, but also forces the old innovators to exit the market [Aghion and Howitt (1992, 1998)].¹⁹

¹⁸(6a) can be derived using the relationships: (i) $\dot{n}_{np}[t] = \dot{n}[t - l]$, (ii) $g_{np}[t] \equiv \frac{\dot{n}_{np}[t]}{n_{np}[t]} = \frac{\dot{n}[t-l]}{n[t-l]} \times \frac{n[t-l]}{n[t]} \times \frac{n[t]}{n_{np}[t]}$, and (iii) $n[t] = n[t - l] e^{\int_{t-l}^{t^-} g[s] ds}$, where the upper bound t^- means that the time is very close to, but not equal to, t . Hence, $g_{np}[t]$ is predetermined by history. (6b) obtains using (1) and (6a).

¹⁹Empirically, in fact, creative destruction does not always mean to destroy old firms. For instance, patented and generic drugs are provided under monopoly and competitive pricing, respectively. Producers of these drugs coexist.

Thus, prior to a patent’s legal maturity, any patented firm faces the hazard of creative destruction at any point in time. Of the n_p patent holders, therefore, n_p^M firms are identified as monopolists and the remaining n_p^{NM} firms are referred to as non-monopolists, who still hold live patents but have lost monopoly power under creative destruction. That is, we have the identity $n_p^M + n_p^{NM} = n_p = n - n_{np}$ at any moment. In fractional terms, this identity is $\zeta_p^M + \zeta_p^{NM} = \zeta_p = 1 - \zeta_{np}$, where $\zeta_p^M \equiv \frac{n_p^M}{n}$ and $\zeta_p^{NM} \equiv \frac{n_p^{NM}}{n}$. The creative-destruction hazard prevails at all times, forcing patented firms to shift from the n_p^M -group to the n_p^{NM} -group over time. This dynamic shift will be formulated later in this subsection.

Hazard rate function in General Let $h[t]$ be the instantaneous hazard-rate at time t and let T be a random variable denoting the time of a patent’s “death” — meaning a patent fails to maintain monopoly power for the patent holder. Then within a patent’s legal life, $h[t]dt$ measures the instantaneous probability that a patented monopolist may unfortunately lose monopoly power in an instant (dt) to newer innovations. In general terms, $h[t]dt$ is a conditional probability defined as²⁰

$$h[t]dt \equiv \Pr[t < T \leq t + dt \mid T > t] = \frac{\Pr[t < T \leq t + dt]}{\Pr[T > t]}, \quad t \in [\tau, \tau + l] \quad (7)$$

where τ denotes the time of a patent issued and $t \in [\tau, \tau + l)$ means that a patent’s legal life is fixed at $l > 0$. In the terminologies of survival analysis, we can define $F[t] \equiv \Pr[T \leq t]$ as the lifetime distribution function, $f[t] = F'[t]$ as the lifetime probability density function, and $S[t] \equiv \Pr[T > t] = 1 - F[t]$ as the survival function. As defined, the survival function measures the probability that a patented firm has survived the ongoing creative-destruction hazards. If the hazard rate $h[t]$ is known at any moment, one can easily derive the survival function as²¹

$$S[t] \equiv \Pr[T > t] = e^{-\Lambda[t]}, \quad \tau \leq t < \tau + l \quad (8)$$

where $\Lambda[t] = \int_{\tau}^t h[\omega]d\omega$ is called the cumulative hazard function for a newly issued patent. This equation implies $S[t] = 1$ at $t = \tau$, connoting that there is no immediate death as the patent is

Apple laptops are pricy at about US\$1,000 or more, but other brands could be as low as \$US500 from anecdotal market data. Apple Inc. has not destroyed all other brands.

²⁰Note that $\Pr[\cdot]$ is a probability function and $\Pr[t < T \leq t + dt \cap T > t] = \Pr[t < T \leq t + dt]$.

²¹From (7) and the definitions of $F[t]$, $f[t]$ and $S[t]$, it holds that $h[t]dt = \frac{f[t]dt}{S[t]} = \frac{-S'[t]}{S[t]}dt = -\frac{d \log S[t]}{dt}dt$. Thus, $h[t] = -\frac{d \log S[t]}{dt}$ and it can solve for (7).

issued. But the patent system is assumed to have a finite patent length, so a terminal boundary condition must be invoked to force $S[t]$ jump down to zero at $t = \tau + l$. This condition forces any surviving patented firm to exit the monopoly group as soon as the patent expires. The hazard rate $h[t]$ and the survival function $S[t]$ play an important role in the model. Later, we will use them to determine the market value of patents of different vintages, and the dynamic changes in patented monopolists and non-monopolists. For now, we need to discuss the market mechanism that determines the time-variant hazard rate $h[t]$ for the modeled economy.

Determining hazard rate In economics literature on innovation and economic growth, economists often deal with multiple sources of hazards that arise from innovating, imitating, or other activities, and they adopt different ways to determine the hazard rate $h[t]$ confronting incumbent firms. The simplest way assumes a constant hazard rate of $h[t] = h_0$ at any time, as in Helpman (1993), Grinols and Lin (1997, 2006), Lin (2002), Furukawa (2007) and others, where the rate of imitation of innovative goods is set exogenously to reflect the degree of intellectual property rights protection and serves as a constant hazard rate. This is equivalent to presuming an exponential lifetime distribution function for the modeled economy so that $F[t] = 1 - e^{-h_0(t-\tau)}$.²² Some other works, however, endogenize the hazard rate either *as an externality of* economy-wide innovating activities [e.g., Grossman and Helpman (1992, ch.4), Jones and Williams (2000), O’Donoghue and Zweimüller (2004), and Iwaisako, et al. (2011)] or *as a purposive result of* profit-driven activities, respectively, in imitation [e.g., Grossman and Helpman (1992, ch. 11), outbound FDI (foreign direct investment) [e.g., Lin (2010)], and rent seeking-cum-innovation [e.g., Dinopoulos and Syropoulos (2007), Grieben and Sener (2009), and Davis and Sener (2012)].²³

Following Jones and Williams (2000), we endogenize the instantaneous hazard rate as an externality of economy-wide innovating activities:

$$h[t] = \psi \cdot \frac{\dot{n}[t]}{n[t]} \equiv \psi \cdot g[t], \quad \psi > 0 \quad (9)$$

²²With this assumed exponential lifetime distribution function, it holds that $f[t] = F'[t] = h_0 e^{-h_0(t-\tau)}$ and $S[t] = e^{-h_0(t-\tau)}$. It must be true that $h[t] = \frac{f[t]}{S[t]} = h_0$ according to (7). It is well-known that the assumed exponential distribution function $F[t] = 1 - e^{-h_0(t-\tau)}$ is associated with a Poisson distribution with probability mass: $\mathcal{P}[x] = \frac{e^{-h_0(t-\tau)} (h_0(t-\tau))^x}{x!}$, where x is the number of event occurrences in a time duration of $t - \tau$ (in most discussion elsewhere, $\tau = 0$ and t represents a time duration). In the context of the present paper, the event refers to creative destruction. It can be verified that $\mathcal{P}[x = 0] = e^{-h_0(t-\tau)} = S[t]$.

²³These cited works all assume an infinite patent length for their quality-ladder growth models.

where $\dot{n} = \frac{dn}{dt}$ as defined earlier. From equation (9), the hazard rate $h[t]$ is positively associated with the innovation rate $g[t]$ determined by the market equilibrium level of aggregate R&D input $L_{R\&D}$ (see (5)). This hazard-rate function becomes lucid if we verify that $h[t]dt$ is actually equal to $\frac{\psi dn}{n} = \frac{\zeta_p^M \cdot \psi dn}{n_p^M}$, a clear measure of the instantaneous probability that a firm among the n_p^M patented monopolists is forced to prematurely lose monopoly power.²⁴

How important is an endogenous hazard rate? First of all, with a non-zero hazard rate, either exogenous or endogenous, effective patent life (i.e. a patented firm's expected monopoly duration) must be shorter than its legal life. More importantly, if the hazard rate is endogenized, a patent's effective patent life must respond to changes in statutory patent length. In this sense, the present paper actually deals with endogenous intellectual property rights (IPR) protection by adjusting statutory patent length to influence the innovation-determined hazard rate, which then determines effective patent life. Therefore, this paper is an important supplement to the endogenous IPR literature that focuses on rent-seeking activities; see, for instance, Dinopoulos and Syropoulos (2007), Eicher and García-Peñalosa (2008), Grieben and Sener (2009), Davis and Sener (2012).

2.1.4 Motions of the fractions of patented monopolists and non-monopolists

We now show how the hazard rate $h[t]$ interact with the innovation rate $g[t]$ so as to drive the shifting dynamics between the n_p^M monopolists and n_p^{NM} non-monopolists. First of all, given the hazard rate $h[t]$, there are $h[t] \cdot n_p^M$ monopolists are forced into the non-monopolist group, where ever firm still holds a legally live patent. As such, it must hold that $\dot{n}_p^{NM}[t] = h[t] \cdot n_p^M[t]$ at any moment. Using this result and differentiating $\zeta_p^M \equiv \frac{n_p^M}{n}$ and $\zeta_p^{NM} \equiv \frac{n_p^{NM}}{n}$ with respect to time t ,

²⁴From equation (9) and $\zeta_p^M = \frac{n_p^M}{n}$, it holds that $h[t]dt = \frac{\psi dn}{n} = \frac{(n_p^M/n) \cdot \psi dn}{n_p^M} = \frac{\zeta_p^M \cdot \psi dn}{n_p^M}$. This results implies that as soon as dn new goods are invented in an instant (dt), there are ψdn existing goods confronting these newly invented goods. Further, assume that the market confrontation spread out proportionally over a stream of all n ($= n_{np} + n_p^{NM} + n_p^M$) existing goods that satisfy the identity $\zeta_p + \zeta_{np} = \zeta_p^M + \zeta_p^{NM} + \zeta_{np} = 1$. Then only $\zeta_p^M \cdot \psi dn$ existing goods, which are monopoly-provided, are forced to lose monopoly power (we call it creative destruction), while all other $(\zeta_{np} + \zeta_p^{NM}) \cdot \psi dn$ goods, which are already competitively provided, merely face more competition in a more crowded differentiated goods market. In contrast to the present paper, Jones and Williams (2000) assume that patents are infinitely lived and that creative destruction is an incidence of forcing incumbent firms to exit the market. Thus, in Jones and Williams (2000), there is no need to distinguish between patented and unpatented firms or between monopolists and non-monopolists.

we can obtain the motions of ζ_p^{NM} and ζ_p^M :²⁵

$$\dot{\zeta}_p^{NM}[t] = \zeta_p^M[t]h[t] - \zeta_p^{NM}[t]g[t] \quad (10a)$$

$$\begin{aligned} \dot{\zeta}_p^M[t] &= g_p[t]\zeta_p[t] - \zeta_p^M[t](h[t] + g[t]) \\ &= (g[t] - g[t-l]e^{-\bar{g}[t]l}) - \zeta_p^M[t](h[t] + g[t]) \end{aligned} \quad (10b)$$

As expected, starting from the model's long-run steady state (i.e. $\dot{\zeta}_p^{NM} = \dot{\zeta}_p^M = 0$), an increase in the hazard rate $h[t]$ must drive the non-monopolist fraction ζ_p^{NM} to increase and the monopolist fraction ζ_p^M to decrease, ceteris paribus. In what follows we turn to the patent value, which determines the value of a patented firm that has not yet been creatively destructed.

2.2 Patent value

In the modeled economy, patents are issued sequentially over time when new varieties are developed. They are tradeable and distinguished by patent vintages (i.e. dates of patent issuance). Define $v[t, \tau]$, $t \geq \tau$, as the value of a patent previously issued at time τ but currently live at time $t \in [\tau, \tau + l)$. Then the patent value $v[t, \tau]$ determines the value of a firm holding a vintage- τ patent. At any point in time, the firm may survive to earn the profit flow $\pi[t, \tau] > 0$, but this instantaneous probability is determined by the survival function $S[\cdot]$ according to (8). The value of a firm holding a vintage- τ patent at time t is therefore determined by the discounted present value of the profit flow it expects to earn during its patent's remaining life ($= \tau + l - t$) to maturity,²⁶

$$\begin{aligned} v[t, \tau] &= \int_t^{\tau+l} e^{-\int_t^s r[\omega]d\omega} \pi[s] \cdot S[s] ds, \quad t \in [\tau, \tau + l) \\ &= \int_t^{\tau+l} e^{-\int_t^s (r[\omega] + h[\omega])d\omega} \pi[s] ds \end{aligned} \quad (11)$$

²⁵Note that $\zeta_p^M + \zeta_p^{NM} = \zeta_p$ and thus $\dot{\zeta}_p^M + \dot{\zeta}_p^{NM} = \dot{\zeta}_p$. Using $\dot{\zeta}_p^{NM}[t] = \left(\frac{\dot{n}_p^{NM}[t]}{n_p^{NM}[t]} - g[t]\right) \zeta_p^{NM}[t] = \left(\frac{n_p^M[t]}{n_p^{NM}[t]} \frac{\dot{n}_p^{NM}[t]}{n_p^M[t]} - g[t]\right) \zeta_p^{NM}[t]$, we obtain (10a) if the equality relationships $\frac{n_p^M}{n_p^{NM}} = \frac{\zeta_p^M}{\zeta_p^{NM}}$ & $\frac{\dot{n}_p^{NM}[t]}{n_p^M[t]} = h[t]$ are used. In similar ways, using $\dot{\zeta}_p^M[t] = \left(\frac{\dot{n}_p^M[t]}{n_p^M[t]} - g[t]\right) \zeta_p^M[t] = \left(\frac{\dot{n}_p[t] - \dot{n}_p^{NM}[t]}{n_p^M[t]} - g[t]\right) \zeta_p^M[t] = \left(\frac{(\dot{n}_p[t]/n_p[t])(n_p[t]/n_p^M[t])}{n_p^M[t]/n_p[t]} - h[t] - g[t]\right) \zeta_p^M[t]$, we obtain (10b).

²⁶Note that the survival function is given by $S[s] = e^{-\Lambda[s]}$, where $\Lambda[s] = \int_t^s h[\omega]d\omega$ compute cumulative hazards from t to s . In equation (8), $\Lambda[t]$ is to compute cumulative hazards from τ to t .

where r is the interest rate, h is the hazard rate, $r + h$ is the risk-adjusted cost of capital, and $\pi[s] > 0$ if $t \in [\tau, \tau + l)$. From (11), we can derive the Bellman-type equation,²⁷

$$\dot{v}[t, \tau] \equiv \frac{dv[t, \tau]}{dt} = r[t]v[t, \tau] - \pi[t] + v[t, \tau] \cdot h[t], \quad t \in [\tau, \tau + l) \quad (12)$$

subject to two boundary conditions

$$v[t, \tau] = \frac{\theta w[t]}{(1 - \phi g[t])n[t]} \quad \text{at } t = \tau \quad (13a)$$

$$v[t, \tau] = 0 \quad \text{at } t \geq \tau + l \quad (13b)$$

(12) is a first-order ordinary differential equation depicting the motion of the value of a live patent and can be easily translated into the well-known no-arbitrage condition — $r[t] = \frac{\dot{v}[t, \tau]}{v[t, \tau]} + \frac{\pi[t]}{v[t, \tau]} - h[t]$ — implying that under well functioning financial markets, the riskless rate of interest r must match the risk-adjusted rate of return (= capital-gain rate $\frac{\dot{v}}{v}$ plus profit rate $\frac{\pi}{v}$ minus hazard rate h). The two boundary conditions, however, constrain the value of any live patent into a closed interval $[0, \frac{\theta w[t]}{(1 - \phi g[t])n[t]}]$, where term $\frac{\theta w[t]}{(1 - \phi g[t])n[t]}$ is the development cost of a new product according to (3) or (5) and represents the upper bound of patent values of all vintages. If $v[t, t]$ were allowed to exceed the upper bound, there would be unbounded R&D demand for labor. It is evident from (11) that $0 \leq v[t, \tau'] < v[t, \tau''] \leq v[t, t] = \frac{\theta w[t]}{(1 - \phi g[t])n[t]}$ for $\tau' < \tau'' \leq t$. That is, consistent with our intuition, a patent with a longer life expectancy is more valuable, given that every monopolist earns the same profit flow at a point in time.

2.3 Households

The household sector provides equity funds to establish firms. We assume a representative household who consumes a stream of goods $\{x_i\}_{i=0}^n$ resulting from previous R&D activities. Hence, total

²⁷Using finite differences, the value of a patent can be approximated by,

$$v[t, \tau] = \pi[t, \tau]\Delta t \cdot (1 - h[t]\Delta t) + \left(\frac{1}{1 + r\Delta t}\right)v[t + \Delta t, \tau] \cdot (1 - h[t]\Delta t)$$

where $h[t]\Delta t$ is the probability of being “creatively destructed” during an instant. Rearranging terms yields

$$\frac{dv[t, \tau]}{dt} \approx \frac{v[t + \Delta t, \tau] - v[t, \tau]}{\Delta t} = rv[t, \tau] - \pi[t, \tau] \cdot (1 - h[t]\Delta t) + v[t + \Delta t, \tau] \cdot h[t]$$

Thus, (12) obtains by letting Δt going to zero asymptotically.

nominal consumption expenditure is equal to $E = \int_0^n p[i]x[i]di$. The household is endowed with a fixed level of labor L and earns the full-employment level of wage income wL , where w is the wage rate. At any point in time, labor L_x is employed to produce goods x_i for $i \in [0, n]$ and the remaining labor $L_{R\&D}$ is engaged in R&D to invent \dot{n} new goods, which result in \dot{n} new patents (or new firms). Therefore, $L = L_x + L_{R\&D}$ is the full employment condition.

The household owns equity shares, which represent claims to all firms. Her total wealth at a time, denoted by A , is the sum of stock market values of all live patents that have not yet been “creatively destroyed” by newer product innovations. That is, $A[t] = \int_\tau^{\tau+l} n_p^M[t, s]v[t, s]ds$, where $n_p^M[t, \tau]$ is the number of valuable vintage- τ patents and satisfies $\int_\tau^{\tau+l} n_p^M[t, s]ds = n_p^M[t]$, the number of monopoly firms. The household earns investment income flows, which are dividends resulting from monopoly profits plus capital gains due to appreciation of share values net of the losses of creative destruction. As such, the net investment income flow must be equal to $r[t]A[t]$ under no-arbitrage conditions.²⁸ The household’s decision problem is to choose $C[t]$ to maximize lifetime utility,

$$U[t] = \int_t^\infty e^{-\rho s} \log\left[\frac{C[s]}{n^{\beta/(\varepsilon-1)}}\right] ds, \quad (14)$$

subject to the dynamic budget constraint,

$$\dot{A}[t] = r[t]A[t] + w[t]L[t] - E[t]. \quad (15)$$

In the utility function parameter ρ reflects the household’s time preference, aggregate consumption is presumably measured by the CES consumption index $C = \left(\int_0^n x[i]^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}}$, parameter $\varepsilon > 1$ is the elasticity of substitution between goods, and term $n^{\beta/(\varepsilon-1)}$ provides flexibility for the utility function to capture the welfare effects of increased variety of substitutable goods. As will be explained later, parameter $\beta \in [0, 1]$ in (14) is used to reflect the aggregate welfare importance of (horizontal) product differentiation.

As (15) indicates, if the aggregate income inflow (net investment income rA plus wage income

²⁸Using the above-mentioned no-arbitrage condition $r[t] = \frac{\dot{v}[t, \tau]}{v[t, \tau]} + \frac{\pi[t]}{v[t, \tau]} - h[t]$, the household’s net investment income flow is given by $A[t] \times \left(\frac{\dot{v}[t, \tau]}{v[t, \tau]} + \frac{\pi[t]}{v[t, \tau]} - h[t]\right)$ and must be tied to $r[t]A[t]$.

wL) exceeds the aggregate consumption expenditure flow E , the household's wealth increases (i.e. $\dot{A} > 0$); otherwise, the reverse is true. Let P be the aggregate price level such that $E = P \times C$. Then, maximization of lifetime utility (14) yields the well-known Euler condition,

$$\frac{\dot{C}[t]}{C[t]} + \frac{\dot{P}[t]}{P[t]} = \frac{\dot{E}[t]}{E[t]} = r[t] - \rho, \quad (16)$$

to govern the instantaneous growth rate $\frac{\dot{E}}{E}$ of nominal consumption expenditure.

For a given level of consumption expenditure E at a point in time, maximization of the instantaneous utility flow measured by $\log[\frac{C[s]}{n^{\beta/(\varepsilon-1)}}]$ yields the goods demand functions,

$$x[i] = \left(\frac{p[i]}{P}\right)^{-\varepsilon} \left(\frac{E}{P}\right), \quad i \in [0, n], \quad (17)$$

where the aggregate price level is given by $P = (\int_0^n p[i]^{1-\varepsilon} di)^{1/(1-\varepsilon)}$ and it can be verified that $E = P \times C$.

2.4 Firms

Assume that a unit of good $x[i]$, $i \in [0, n]$, is produced with a unit of labor. Then the marginal cost of any available good is equal to labor wage w . As described above, firms with dead patents are perfectly competitive, so are the firms that hold legally valid patents but had been forced by creative destruction to give away monopoly power to newer innovative firms. The two groups of firms therefore price outputs at marginal cost. The third group of firms hold legally live patents and are still able to command their patent-granted monopoly power, so they price outputs at a markup over marginal cost. Assume that each group of firms are symmetric. Their pricing conditions are given below:

$$p[i] = \begin{cases} p_{np} = w & \text{if } i \in [0, n_{np}] \\ p_p^{NM} = w & \text{if } i \in NM \subset (n_{np}, n] \\ p_p^M = \alpha w & \text{if } i \in M \subset (n_{np}, n] \end{cases} \quad (18)$$

where p_{np} is the representative price of goods from firms with dead patents, p_p^{NM} is the representative price of goods from firms with legally live patents but with no monopoly power, and p_p^M is the representative price of goods from firms with both legally live patents and monopoly power.

Both p_{np} and p_p^{NM} are equal to marginal cost w , which is below the monopoly price p_p^M , subject to the markup rate $\alpha = \frac{\varepsilon}{\varepsilon-1} > 1$. These representative prices and demand functions (17) imply the following representative outputs in goods market equilibrium:

$$x[i] = \begin{cases} x_{np} = \left(\frac{p_{np}}{P}\right)^{-\varepsilon} \left(\frac{E}{P}\right), & \text{if } i \in [0, n_{np}], \\ x_p^{NM} = \left(\frac{p_p^{NM}}{P}\right)^{-\varepsilon} \left(\frac{E}{P}\right), & \text{if } i \in NM \subset (n_{np}, n], \\ x_p^M = \left(\frac{p_p^M}{P}\right)^{-\varepsilon} \left(\frac{E}{P}\right), & \text{if } i \in M \subset (n_{np}, n]. \end{cases} \quad (19)$$

Clearly, $x_{np} = x_p^{NM} = \alpha^\varepsilon x_p^M$ and using these representative outputs, price level P can simplify to

$$P = (n_{np} p_{np}^{1-\varepsilon} + n_p^{NM} (p_p^{NM})^{1-\varepsilon} + n_p^M (p_p^M)^{1-\varepsilon})^{\frac{1}{1-\varepsilon}} = w \cdot n^{-\frac{1}{\varepsilon-1}} \cdot \varphi, \quad (20a)$$

where

$$\varphi = \varphi[\zeta_p^M, \alpha] = \left(1 - \zeta_p^M + \zeta_p^M \alpha^{1-\varepsilon}\right)^{1/(1-\varepsilon)} \geq 1 \quad (20b)$$

As indicated, $\varphi \geq 1$ is a measure of monopoly distortion, raising the aggregate price level P . This distortion is significant if φ deviates further from one, and it vanishes if φ falls to one. From (18) and (19), every monopolist earns a positive profit flow equal to $\pi[t, \tau] = \pi[t] = (\alpha - 1)w x_p^M$ at a time if $t \in [\tau, \tau + l)$, regardless of the vintage of the patent. However, the profit flow $\pi[t]$ per firm will diminish over time as future product innovation will decrease the market size x_p^M of every incumbent monopolist. This will become clear later in subsection 2.5.

Now, substituting (19) into (14) permits us to write the log flow of utility as

$$\begin{aligned} \log\left[\frac{C}{n^{\beta/(\varepsilon-1)}}\right] &= \log\left[n^{\frac{1}{\varepsilon-1}(1-\beta)}\right] \\ &+ \log\left[\left((\zeta_{np})^{\frac{1}{\varepsilon}} (X_{np})^{\frac{\varepsilon-1}{\varepsilon}} + (\zeta_p^{NM})^{\frac{1}{\varepsilon}} (X_p^{NM})^{\frac{\varepsilon-1}{\varepsilon}} + (\zeta_p^M)^{\frac{1}{\varepsilon}} (X_p^M)^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}\right] \end{aligned} \quad (21)$$

where $X_{np} = n_{np} x_{np}$, $X_p^{NM} = n_p^{NM} x_p^{NM}$ and $X_p^M = n_p^M x_p^M$ are aggregate consumption of goods from each group of firms, respectively. From (21), we can see how parameter β affects the log-flow of utility. If β is between 0 and 1, an increase in n contributes to the utility flow, but at a decreasing rate due to the logarithmic term. If $\beta = 1$, an increase in n has no effect on utility, ceteris paribus. Hence, parameter β reflects the aggregate welfare importance of product differentiation. We assume

$\beta \in [0, 1]$ in the model, in recognition that greater prevalence of similar goods exerts diminishing welfare effects and that differentiated goods are often not equally important to regional demanders in the world.²⁹

2.5 Full employment and economy-wide markup

We now proceed to inspect how product innovation can dilute the profit flow over time. As noted earlier, R&D employment of labor is equal to $L_{R\&D} = \frac{\theta g}{1-\phi g}$ in terms of (3), so $L - L_{R\&D}$ is labor employed to manufacture goods. The full employment condition can therefore be rewritten as

$$L - \frac{\theta g}{1 - \phi g} = n \cdot (\zeta_p^M x_p^M + \zeta_p^{NM} x_p^{NM} + \zeta_{np} x_{np}). \quad (22)$$

Using $x_{np} = x_p^{NM} = \alpha^\varepsilon x_p^M$ and $\zeta_p^M + \zeta_p^{NM} + \zeta_{np} = 1$, the levels of outputs at full employment can be derived as follows:

$$x_{np} = x_p^{NM} = \frac{1}{n} \left(\frac{\alpha^\varepsilon}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \quad (23a)$$

$$x_p^M = \frac{1}{n} \left(\frac{1}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \quad (23b)$$

It is now clear from (23a) and (23b) that given a constant innovation rate g , the growth of product variety (greater n) must decrease the market size for all incumbent firms. Thus, as mentioned in (11), the profit flow $\pi[t]$ accrued to every monopolist must be diluted over time by future product innovation. Using (23a) and (23b), we can further pin down aggregate consumption expenditure $E = n(1 - \zeta_p^M)x_{np} \cdot w + n\zeta_p^M x_p^M \cdot \alpha w$ as shown below,

$$E = \xi w \left(L - \frac{\theta g}{1 - \phi g} \right), \quad (24)$$

where

$$\xi = \xi[\zeta_p^M] = \frac{\zeta_p^M \alpha + (1 - \zeta_p^M)\alpha^\varepsilon}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon} < \alpha \quad (25)$$

Variable ξ bears interesting economic implications. From (24), ξ is the ratio of consumption expenditure E to aggregate production cost $w(L - \frac{\theta g}{1-\phi g})$ of all available goods. Thus, we can inter-

²⁹See Footnote 8 for more detailed discussions.

pret ξ as an economy-wide markup rate, which is endogenous and varies with the fraction ζ_p^M of monopoly goods, given the exogenous product-specific markup rate α . It can be shown that (i) the economy-wide markup rate is positively associated with the fraction of monopoly goods due to $\frac{\partial \xi}{\partial \zeta_p^M} = \frac{(\alpha-1)\alpha^\varepsilon}{[\zeta_p^M + (1-\zeta_p^M)\alpha^\varepsilon]^2} > 0$ for $\zeta_p^M \in (0, 1)$; (ii) $\xi = \alpha$ at $\zeta_p^M = 1$ implying that all available goods are priced at a markup over marginal cost; (iii) $\xi = 1$ at $\zeta_p^M = 0$ implying that all goods are priced at marginal cost.

2.6 Balanced growth path

We chose E to be a numeraire; i.e., $E = 1$. Then given a set of state variables $n, n_p, n_{np}, n_p^{NM}, n_p^M, \zeta_p, \zeta_{np}, \zeta_p^{NM}, \zeta_p^M$, which are predetermined at any point in time, we can solve the model's nine instantaneous equilibrium conditions (18), (19), (20a), (22), (24) for the nine variables $p_p^M, p_p^{NM}, p_{np}, w, x_p^M, x_p^{NM}, x_{np}, P$, and g .³⁰ As the time marches forward, however, the predetermined state variables of n 's will keep growing, provided the economy has ongoing product innovation, and the other predetermined state variables of ζ 's will change in transition toward their respective long-run steady-state equilibrium levels. As the paper is focused on the question of what patent length (l) maximizes a global economy's long-run welfare, we confine the analysis to the model's balanced growth path in what follows.

Hence, we need a set of long-run steady state equilibrium conditions to close the model. First, choosing nominal expenditure E as the numeraire ensures that $\dot{E}[t] = 0$ and thus $r[t] = r = \rho$ at all times according to the Euler condition (16). Second, a balanced growth path requires that the innovation rate $g[t]$ must stay constant over time so that $g[t-l] = g[t] = g$ at any moment. Third, along the balanced growth path, the fractions of ζ 's must stay constant as well; that is, $\dot{\zeta}_{np}[t] = \dot{\zeta}_p[t] = \dot{\zeta}_p^{NM}[t] = \dot{\zeta}_p^M[t] = 0$ at all times. Fourth, we applied the patent value's upper bound (13a) to the no-arbitrage condition (11) in order to ensure an ongoing innovation process in steady state. In so doing, we derived below a two-equation, steady-state equilibrium system that

³⁰Once g is solved, h is determined in terms of (9), and g_p is determined in terms of (6b), given that g_{np} is predetermined by (6a).

can be solved for the values of g and ζ_p^M along a balanced growth path:³¹

$$\mu = (\alpha - 1) \cdot \left(\frac{1}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon} \right) (L - \mu g) \left(\frac{1 - e^{-(\rho+h+g)l}}{\rho + h + g} \right) \quad (26a)$$

$$\zeta_p^M = \frac{1 - e^{-gl}}{1 + \psi} \quad (26b)$$

where $\mu = \frac{\theta}{1-\phi g}$ is an inverse measure of research productivity from equation (4) and $h = \psi g$ is the instantaneous hazard rate according to equation (9). Remarks on each of the two steady state equilibrium conditions are in order. For condition (26a), the very left term of μ reflects the real cost (measured in labor units) of developing a new product, while the right-hand side measures the discounted present value (measured in labor units) of future profit flows accruing to a newly established monopoly firm.³² On the right-hand side are four interesting terms: (i) term $\alpha - 1$ is the profit flow per output unit; (ii) term $\frac{1}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon}$ is the market share per monopolist; (iii) term $L - \mu g$ is the combined output level of available goods,³³ and (iv) term $\frac{1 - e^{-(\rho+h+g)l}}{\rho+h+g}$ is the effective discount factor, implying that the “effective cost of capital” actually equals $\rho + h + g (= r + h + g)$ along the balanced growth path. The innovation rate g enters the effective discount factor to capture innovation-induced profit dilution. Another important aspect is that if patent length l were set equal zero, the effective discount factor would become zero, thereby making a viable innovating process impossible. But as we will show later, the minimum patent length to sustain a viable innovating process is actually way above zero. The next steady-state condition (26b) determines the loci of (ζ_p^M, g) that ensures $\dot{\zeta}_p^M = 0$, given a constant innovation rate g . This condition implies that a greater constant innovation rate must be associated with a greater constant fraction ζ_p^M of monopoly firms in steady state.

If we let g approach zero in conditions (26a) and (26b), the minimum patent length l required

³¹Derivations of (26a) are shown in Appendix A, while (26b) can be easily derived by setting $\dot{\zeta}_p^M = 0$ in (10b) where $\bar{g}[t] = \frac{1}{l} \int_{t-l}^t g[s] ds = g$ in steady state. Notice that all other state variables can be expressed in terms of g and ζ_p^M . For instance, noting $g[t] = g[t - l] = g$ in steady state, we have $\zeta_{np} = e^{-gl}$ from (6a), $\zeta_p = 1 - e^{-gl}$ from (6b), and $\zeta_p^{NM} = \zeta_p - \zeta_p^M$.

³²It should be noticed that the measure of product variety, n , drops out of condition (26a). As analyzed earlier, given a constant innovation rate g , the growth of product variety n must dilute per firm profit flow over time. But on the other hand, the growth of product variety (or stock of knowledge) also keeps decreasing the development cost due to the positive externality of knowledge stock. The two opposing forces precisely counter each other so that product variety n does not appear in (26a). See Appendix A for details.

³³The product of μ and g measures the level of labor employment in R&D activities. So, the remaining labor $L - \mu g$ is employed to manufacture goods.

to sustain a viable innovation process in steady state is given by

$$\underline{l} = \frac{-\log \left[1 - \left(\frac{\alpha^\varepsilon}{\alpha-1} \right) \left(\frac{\theta}{L} \right) \rho \right]}{\rho} \quad (27)$$

As l approaches \underline{l} from above, both g and ζ_p^M go to 0^+ . As noted earlier, the R&D production function (5) implies that an active R&D sector requires a minimum patent length. Thus far we have developed a model ready for computational analysis.

3 Model Calibrating and Simulations

3.1 Baseline parameter values

To calibrate the model to a closed economy, which we take as a global economy, we chose a set of baseline parameter values in Table 1.³⁴ The coefficient ρ of time preference was set equal to 0.04, which ensures the real interest rate r equal to 4 percent per annum, since E serves as a numeraire. The elasticity ε of demand for an individual product was set equal to 1.23, which ensures an individual monopoly product's markup rate of $\alpha = \varepsilon/(\varepsilon - 1) = 5.3$. Assuming a globally harmonious patent system, patent length l was chosen as 20 years to match the level of advanced economies in the world. $L = 1$ implies that global population was normalized to unity. Finally, we are free to choose the other three parameter values (θ , ϕ , ψ) on account of no reliable empirical estimates. We therefore chose $\theta = 5$ (inverse of research productivity), $\phi = 1$ (research congestion or patent race) and $\psi = 1$ (creative destruction), so as to make the model able to deliver the innovation rate (g), economy-wide markup rate (ξ), cost of capital ($\rho + h$) and effective cost of capital ($\rho + h + g$) that largely mimic the world's empirical statistics.

Parameter	ρ	ε	l	L	θ	ϕ	ψ
Value	0.04	1.23	20	1	5	1	1

Table 1: Baseline parameter values

³⁴The chosen parameter ρ (time preference) is in annual rates so that the time unit in the simulations represents one year. As such, $l = 10$ means a 10-year patent length.

3.2 Computing equilibrium along balanced-growth path

Using the baseline parameter values listed in Table 1 and for illustrative purposes, we computed the steady-state equilibrium by solving equations (26a) and (26b), respectively, for the patent length of $l = 20$ years and $l \rightarrow \infty$. The results are shown in Figure 1, where point *a* represents the benchmark steady state equilibrium with the fraction of monopoly firms $\zeta_p^M = 21.67\%$ and the innovation rate $g = 2.84\%$. Using the computed innovation rate and (9), the creative-destruction hazard rate is equal to $h = 0.0284$, which implies the cost of capital equal to $\rho + h = 0.0684$ and the effective cost of capital equal to $\rho + h + g = 0.0968$. The computed monopoly fraction $\zeta_p^M = 21.67\%$ and the chosen demand elasticity $\varepsilon = 1.23$ combine to yield the economy-wide markup rate of $\xi = 1.15$ according to (25). The economy-wide markup rate falls onto the range of estimates 1.05 to 1.40 by Norrbin (1993) and Basu (1996). The innovation rate of 2.84% also largely matches the average annual growth rate of the world's OECD-published triadic patent families between 1985 and 2008.

As indicated in Figure 1, if patent length l is infinitely large, the locus of (ζ_p^M, g) at which $\dot{\zeta}_p^M = 0$ shift rightwards from the solid curve to the dashed one, while the locus at which $\dot{g} = 0$ shifts upwards from the solid curve to the dashed one. As a result, the steady state equilibrium shifts from point *a* to point *b*, which indicates a faster innovation rate ($g = 4.92\%$) and a larger monopoly fraction ($\zeta_p^M = 50\%$). These results yield a greater hazard rate ($h = 0.0492$), a greater cost of capital ($\rho + h = 0.0892$), a greater effective cost of capital ($\rho + h + g = 0.1384$), and a greater economy-wide markup rate of $\xi = 1.49$. These numerical results reflect the present model's novel features. First of all, as l is lengthened to unlimited years, larger future profit flows redirect more resources (labor) into R&D activities, thereby generating a faster innovation rate (from 2.84% to 4.92%). But the faster innovation rate must give rise to a larger idiosyncratic risk of creative destruction on incumbent monopolists, thereby raising the hazard rate and the cost of capital. Also, faster innovation must raise the weight of monopoly and thus the economy-wide markup rate. Further, as formulated in (4), due to the prevalence of patent races, a faster innovation rate can lead to lower research productivity at a given stock of knowledge (denoted by n). Without the hazard of creative destruction and/or the prevalence of patent races, however, the innovating pace resulting from increased patent life would otherwise be even stronger, as we will report in what follows.

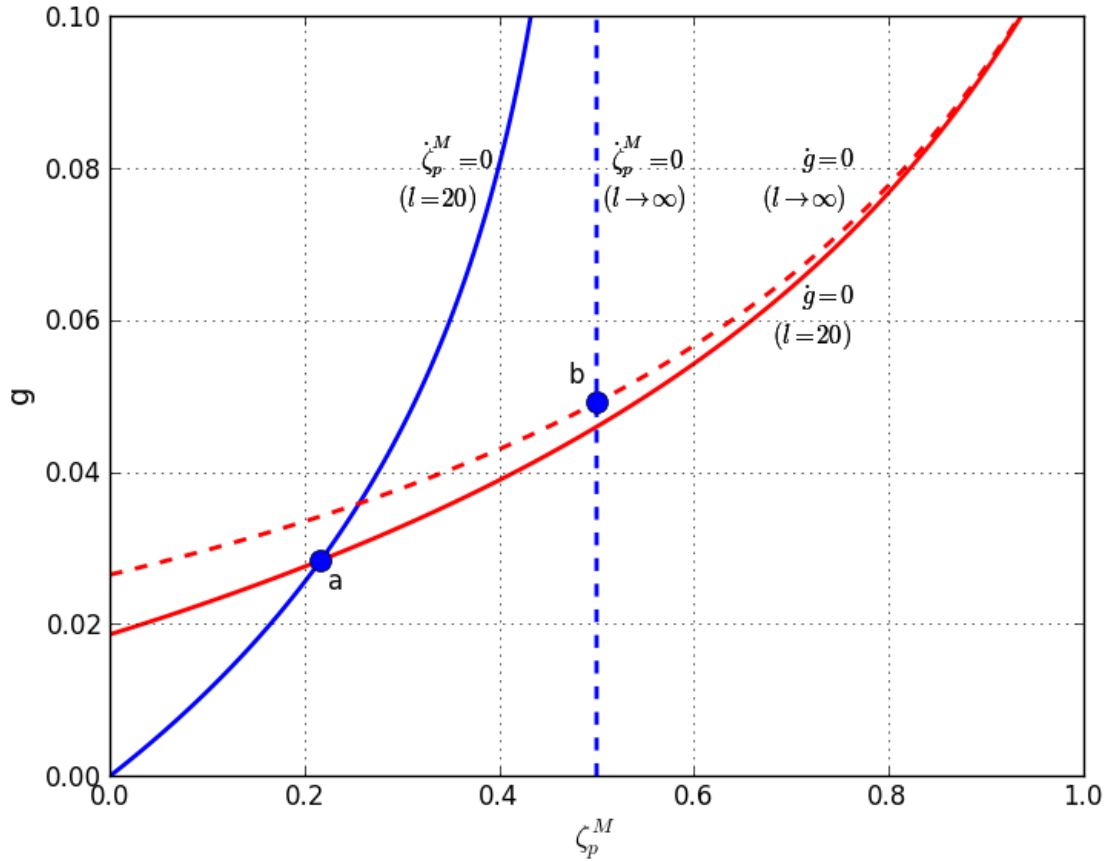


Figure 1: Steady State Equilibrium at $l = 20$ or $l \rightarrow \infty$, using baseline parameter values for eqs. (26a) & (26b).

3.3 Robustness checks on patent length and pace of innovation

From the above simulations, extending the benchmark patent length (20 years) into infinity was seen to enhance the innovation rate, the fraction of monopoly goods, and the hazard rate of creative destruction. But we are yet to see whether these relationships still hold when only smaller changes in patent length l are in place. In this subsection we simulate the effects of extending the patent length. Scenario A of Table 2 represents the benchmark scenario using the baseline parameter values. The implications of Scenario A is summarized in Proposition 1:

Proposition 1: For a closed economy, extending the duration of patents tends to increase the innovation rate, the share of monopoly goods, the economy-wide markup, and the hazard rate facing incumbent monopoly firms along the balanced growth path,

Patent Term l	Scenario A $\psi = 1, \phi = 1$			Scenario B: $\psi \rightarrow 1.05$			Scenario C: $\phi \rightarrow 1.05$		
	$g, \%$	$\zeta_P^M, \%$	h	$g, \%$	$\zeta_P^M, \%$	h	$g, \%$	$\zeta_P^M, \%$	h
12	0.44	2.58	.0044	0.43	2.44	.0045	0.44	2.57	.0044
15	1.68	11.15	.0168	1.63	10.55	.0171	1.67	11.11	.0167
20	2.84	21.67	.0284	2.74	20.58	.0288	2.83	21.60	.0283
25	3.48	29.06	.0348	3.36	27.70	.0352	3.47	28.99	.0348
30	3.88	34.40	.0388	3.74	32.90	.0393	3.87	34.33	.0387
50	4.59	44.95	.0459	4.42	43.43	.0464	4.57	44.91	.0457
100	4.89	49.62	.0489	4.73	48.35	.0497	4.88	49.62	.0488
∞	4.92	50.00	.0492	4.76	48.78	.050	4.90	50.00	.0490

Table 2: Scenario A is benchmark with $\psi = \phi = 1$. Product innovation, monopoly firms, and hazard rate, as ψ (coefficient of creative destruction) or ϕ (coefficient of research congestion or patent races) is increased.

but these effects tend to diminish with the patent length.³⁵

These diminishing marginal effects of extending the patent term appear robust to changes in the value of parameters ψ and ϕ , which, as noted earlier, respectively, capture the effects of creative destruction and research congestion/patent race. The robustness checks are presented, respectively, in Scenario B of Table 2. From these robustness checks, we obtain:

Proposition 2: For a closed economy where product innovation creates the hazard of creative destruction and the wasting of resources via research congestion/patent races, the innovation rate corresponding to each patent term tends to be smaller along the balanced growth path, if either the effect of creative destruction or the prevalence of research congestion or patent race becomes more significant.³⁶

4 Welfare

We now turn to welfare analysis and demonstrate how to numerically solve for the socially optimal patent length.

³⁵The diminishing effects of extending the patent term appear clearly from Scenario A of Table 2. For instance, as patent term l is increased from 15 to 20 years (an increase of 5 years), the innovation rate g is up 1.16 percentage points ($= 2.84\% - 1.68\%$), but as l is increased from 20 to 25 years (an increase of 10 years), the innovation rate is only up 0.64 percentage point ($= 3.48\% - 2.84\%$).

³⁶For instance, at $l = 20$, the innovation rate is $g = 2.74\%$ from scenario B (larger ψ) and $g = 2.83\%$ from scenario C (larger ϕ), compared to $g = 2.84\%$ from benchmark scenario A.

4.1 Channels of welfare change

For welfare analysis, we decomposed lifetime utility (14) into a predetermined constant term $\frac{\log n[0]}{(\varepsilon-1)\rho}$, which can be suppressed by setting $n[0] = 1$, plus four endogenous components attributed to product innovation (U^n), saving rate (U^s), monopoly-induced price distortion ($U^{M/price}$), and monopoly-induced cost reduction ($U^{M/cost}$) as given below:³⁷

$$U = U^n + U^s + U^{M/Price} + U^{M/Cost} \quad (28)$$

where

$$U^n = \frac{(1-\beta)g}{(\varepsilon-1)\rho^2} \geq 0, \quad (29a)$$

$$U^s = \frac{1}{\rho} \log(L - \mu g) \leq 0, \quad (29b)$$

$$U^{M/price} = -\frac{1}{\rho} \log \varphi = \frac{1}{(\varepsilon-1)\rho} \log(1 - \zeta_p^M + \zeta_p^M \alpha^{1-\varepsilon}) \leq 0, \quad (29c)$$

$$U^{M/cost} = \frac{1}{\rho} \log \xi = \frac{1}{\rho} \log \frac{\zeta_p^M \alpha + (1 - \zeta_p^M) \alpha^\varepsilon}{\zeta_p^M + (1 - \zeta_p^M) \alpha^\varepsilon} \geq 0. \quad (29d)$$

Explanation of these four welfare components are in order. First, with $0 \leq \beta \leq 1$, $\varepsilon > 1$ and $\rho > 0$, the innovation-generated welfare component, U^n , is nonnegative in terms of (29a), provided R&D is sustainable. This component contributes more to lifetime utility if product innovation proceeds at a faster rate. Second, a faster innovation rate (increase in g) must result in a higher degree of R&D intensity measured by μg ($= \frac{\theta g}{1-\phi g}$; see (5)) call for a higher saving rate, and thus decrease lifetime utility.³⁸ That is, faster innovation requires more R&D investment, which calls for more saving from households. Therefore, to sustain a faster pace of product innovation, households must have a higher saving rate to finance a higher research-intensive economy, and thus sacrifice welfare to some extent. This welfare channel is captured by the second (nonpositive) component $U^s \leq 0$ according to (29b). Third, from (26b), a faster innovation rate must raise the monopoly fraction ζ_p^M , thereby deteriorating the monopoly-induced price distortion by raising the aggregate price level P through an increase in $\varphi \geq 1$ and thereby decreasing household welfare, as indicated by the

³⁷See Appendix B for the derivations of lifetime utility components.

³⁸Note that since μg measures the level of R&D labor employment, $L = 1$ implies $(\frac{\theta g}{1-\phi g})/L = \frac{\theta g}{1-\phi g}$, a share of resources (labor) allocated to R&D. So, μg is also a measure of the modeled economy's R&D intensity.

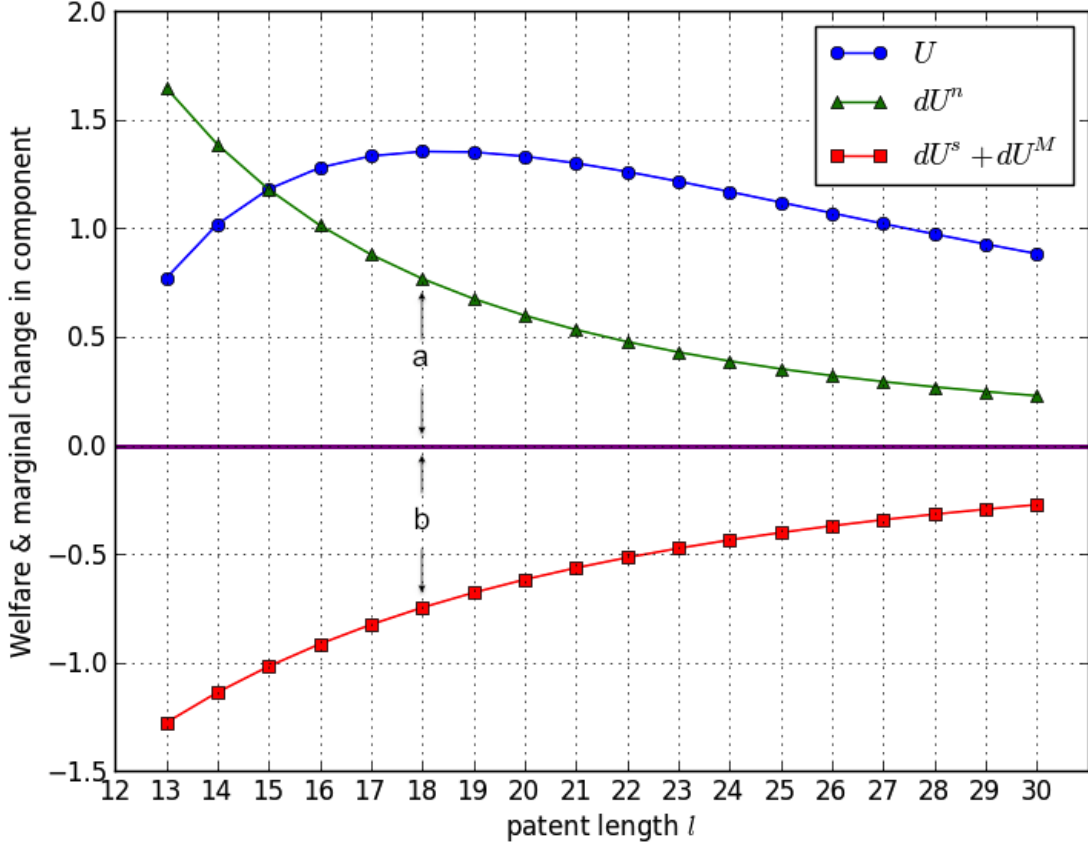


Figure 2: Optimal patent length: balancing gains from innovation against welfare losses from both increased R&D intensity (saving) and increased net monopoly distortion; distances a and b are equal approximately, meaning $dU^n \approx |dU^s + dU^M|$ with $dU^M = dU^{M/price} + dU^{M/cost}$ (given $dl \approx \Delta l = 1$), where $dU(l = i) \approx U(l = i) - U(l = i - 1)$.

third (nonpositive) component $U^{M/price} \leq 0$.³⁹ Lastly, notice that while the monopoly distortion raises P , the same distortion also decreases demand for labor in production of goods. This effect tends to lower labor wage and thus marginal production cost, thereby to some extent mitigating the monopoly-induced price distortion. Hence, the monopoly-induced cost reduction has a nonnegative contribution to household welfare, as indicated by the fourth (nonnegative) component $U^{M/cost} \geq 0$. Thus, in the present general-equilibrium setting, the net monopoly distortion is measured by $U^M \equiv U^{M/price} + U^{M/cost}$.

³⁹From (20b), $\varphi \geq 1$ and therefore $U^{Monopoly/Price}$ must be nonpositive.

4.2 Simulations for optimal patent life

From our earlier simulations in subsection 3.2, a longer duration of the patent life leads to a faster pace of product innovation along the balanced-growth path. But from (28) and (29a) - (29d), extending the patent life l exerts opposing steady-state effects on household's lifetime welfare through multiple channels. Hence, if there exists a finite patent length $l^* \in (L, \infty) \subset [0, \infty)$ that maximizes the household's lifetime steady-state welfare, such an optimal patent life must be determined at the margin where⁴⁰

$$\frac{dU}{dl} = \frac{dU^n}{dl} + \frac{dU^s}{dl} + \frac{dU^M}{dl} = 0.$$

That is, the optimal patent length l^* must ensure that the welfare increment attributed to faster product innovation be balanced against the welfare reduction attributed to both increased research intensity and increased net monopoly distortion. For illustrative purposes, we discretized the l -domain $[0, \infty)$ into a set of equally distanced grid points so that $dl \approx \Delta l = 1$ (i.e., an one-year increment in patent length l). At each grid point $i \in [1, 2, \dots, \infty)$, we computed $U[l = i]$ and $dU^z[l = i] \equiv U^z[l = i] - U^z[l = i - 1]$, $z \in \{n, s, M\}$ using the same baseline parameter values, as in Table 1, plus another baseline β parameter value, which we chose to be 0.875 in order to reflect the aggregate welfare importance of product differentiation.⁴¹ We reported the computational results in Figure 2.

Figure 2 present tradeoffs between the marginal welfare gain ($= dU^n$) and the marginal welfare loss ($= dU^s + dU^M$) for the patent term running from 13 to 30 years. As indicated, the optimal patent length is about equal to 18 years at which extending the patent term results in a marginal gain (measured by distance a) from faster innovation approximately offset by the marginal

⁴⁰Note that $\frac{dU^M}{dl} = \frac{dU^{M/price}}{dl} + \frac{dU^{M/ocst}}{dl}$.

⁴¹To make sense of the quantitative implication of the β parameter, consider (14), which can be interpreted as the discounted present value of the log-flow of "effective" aggregate consumption $\tilde{C} \equiv C/n^{\beta/(\varepsilon-1)}$. In general, we can use $\ln \tilde{C}_1 - \ln \tilde{C}_0$ as an approximation to the percent change in \tilde{C} when the economy switches from scenario 0 to scenario 1. Having noticed this and from (29a), we can interpret $dU^n \approx U_1^n - U_0^n = \frac{(1-\beta)(g_1-g_0)}{(\varepsilon-1)\rho^2}$ as the percent change in the discounted present value of the "effective" aggregate consumption that is attributed to product innovation, when the innovation rate is changed from g_0 to g_1 . Suppose that $g_1 - g_0 = 0.01$, given the baseline parameter values ($\varepsilon = 1.2$, $\rho = 0.04$, $\beta = 0.875$). Then from (29a), the one-percentage point increase in the innovation rate will result in a 391% increase in the "effective" aggregate consumption that is attributed to product innovation in present-value terms. If β were reduced to 0.5, the resulting percent increase in the innovation-induced effective aggregate consumption would sharply raise to 1,563%. If β were further reduced to 0.1, the percent increase would explode to 2,813%. To begin with, we therefore consider $\beta = 0.875$ more empirically plausible than $\beta = 0.5$ or 0.1.

Patent Length	$\psi = 0.80$	$\psi = 0.90$	$\psi = 1$ (benchmark)	$\psi = 1.10$	$\psi = 1.20$
\underline{l}	11.23	11.23	11.23	11.23	11.23
l^*	17.96	18.13	18.35	18.60	18.85

(a) Effects of changes in ψ (coefficient of creative destruction): as ψ is increased, the world calls for a longer optimal patent length, but \underline{l} stays intact.

Patent Length	$\phi = 0.80$	$\phi = 0.90$	$\phi = 1$ (benchmark)	$\phi = 1.10$	$\phi = 1.20$
\underline{l}	11.23	11.23	11.23	11.23	11.23
l^*	18.64	18.49	18.35	18.22	18.10

(b) Effects of changes in ϕ (coefficient of research congestion/patent race): as ϕ is increased, the world calls for a shorter optimal patent length, but \underline{l} stays intact.

Table 3: Sensitivity analysis: optimal patent length and changes in ψ , ϕ .

loss (measured by distance b) from both increased research intensity and increased net monopoly distortion. Note that both the marginal gain and marginal loss of extending the patent life are seen to diminish with the patent length. This is exactly consistent with Proposition 1 of Subsection 3.3

4.3 Golden section search and sensitivity analysis

The above marginal analysis was conducted on a discrete domain in l . This is good for illustrative purposes. Henceforth, we switch to a continuous domain in l in order to see whether or not the optimal patent length is sensitive to the chosen values of some parameters of interest. To this end, we use an algorithm known as “Golden Section Search” to compute the optimal patent length $l^* \in (\underline{l}, \infty)$ that maximizes the objective function $U[l] = U[g[l], \zeta_p^M[l]]$ given by (28), subject to the steady-state system defined by (26a) and (26b) along with the lower bound \underline{l} determined by (27).

The golden section search algorithm, coded with the open-source Python programming language, is introduced in Kiusalaas (2010, pp. 376-379). This algorithm involves the “bracketing” of the optimal patent term l^* in an closed interval like $[a, b] \subseteq (\underline{l}, \infty)$ and the so-called “golden section telescoping” to iteratively narrow the closed interval that bracket l^* until the length of the bracketing interval is reduced to an error tolerance.⁴² Using the golden section search, we found

⁴²Since the patent term l does not directly enter the objective function, I extended Kiusalaas’s Python code to fit in my model’s constrained maximization problem. The extended Python code is available upon request.

Patent Length	$\varepsilon = 1.13$	$\varepsilon = 1.18$	$\varepsilon = 1.23$ (benchmark)	$\varepsilon = 1.28$	$\varepsilon = 1.33$
\underline{l}	8.89	10.05	11.23	12.43	13.68
l^*	1070.00	26.77	18.35	15.52	13.83

(a) Effects of changes in ε (individual importance of product differentiation): as ε is increased (i.e. smaller individual importance), an active R&D sector requires a longer minimum patent length, but the world calls for a shorter optimal patent length.

Patent Length	$\beta = .845$	$\beta = .865$	$\beta = .875$ (benchmark)	$\beta = .885$	$\beta = .905$
\underline{l}	11.23	11.23	11.23	11.23	11.23
l^*	1070.45	21.38	18.35	15.96	12.16

(b) Effects of changes in β (inverse of aggregate importance of product differentiation): as β is increased, the world calls for a smaller optimal patent length, but \underline{l} stays intact.

Table 4: Sensitivity analysis: optimal patent length and changes in ε and β

that the world's optimal patent length equals 18.35 years (actually, not far away from the 18 years term approximated earlier on a discrete domain on l) under the baseline parameter values; see Table 1. We conducted three sets of sensitivity analysis to see how the optimal patent length l^* will respond to changes in some parameter values:

(i) The first set of sensitivity analysis looks at the two parameters (ψ, ϕ) that affect the hazard of creative destruction and the prevalence of research congestion/patent races. The results are summarized in Table 3 including panels 3a and 3b. Inspection of these results implies that the optimal patent length for the world may range from 17 to 19 years. We see that the hazard of creative destruction tends to lengthen the optimal patent term, but to the contrary, but the prevalence of patent races tends to shorten it. These interesting findings make intuitive sense: the hazard of creative destruction reduces a patented firm's expected monopoly duration and therefore calls for a longer optimal patent term, while the prevalence of research congestion/patent races raises the social cost of increased research intensity and therefore calls for a shorter optimal patent term.

(ii) However, from the second set of sensitivity analysis, summarized in Table 4, the optimal patent length appears sensitive to both ε (elasticity of substitution between goods) and β (aggregate welfare importance of product differentiation). From Panel 4a, if goods are more substitutable (i.e. larger ε), the world calls for a shorter optimal patent length. This is consistent with our intuition: too-similar goods do not deserve a longer patent life. From Panel 4b, if the world's aggregate

Patent Length	$\theta = 4.8$	$\theta = 4.9$	$\theta = 5$ (benchmark)	$\theta = 5.1$	$\theta = 5.2$
\underline{l}	10.67	10.94	11.23	11.51	11.80
l^*	18.96	18.63	18.35	18.11	17.89

Table 5: Sensitivity analysis: optimal patent length and changes in θ (inverse of research productivity). As θ is increased, an active R&D sector requires a longer minimum patent length, but the world calls for a shorter optimal patent length.

welfare appreciates varieties of goods in a way strong enough (i.e. smaller β), the optimal patent term can surprisingly extend beyond even 1,000 years! For instance, at $\beta = 0.845$, l^* is as great as 1070.45 years. To the contrary, if innovative goods are only tailored to some affluent regions (say, developed countries), then the world's aggregate welfare may not appreciate varieties of goods as much (i.e. larger β).⁴³ For instance, if β is as large as 0.905, the world may call for an optimal patent length as low as 12.16 years.

(iii) The third set of sensitivity analysis, summarized in Table 5 examines the role of research productivity parameter θ . From equation (4), this parameter, along with ϕ and g , determines the value of $\mu = \theta/(1 - \phi g)$, an inverse measure of R&D productivity, or equivalently, a measure of R&D cost (see (5) & (26a)). Also, recall that μg determines the modeled economy's R&D intensity, with population normalized to unity (see footnote 38). As expected, if labor is less productive in innovating activities (i.e. larger θ), the world requires a longer minimum patent length to keep an active R&D sector. However, a larger value of θ is also meant to raise the economy's R&D intensity, *ceteris paribus*, thereby decreasing steady-state welfare through the saving channel according to (29b) and therefore calling for a shorter optimal patent length to counter the force to raise R&D intensity for the modeled economy.

5 Concluding Remarks

The paper has used a R&D-based endogenous growth model, where product innovation makes more horizontally differentiated goods available to consumers, in order to study the relationships between patent length, pace of innovation, and aggregate welfare. The model was calibrated to a global economy with a population normalized to unity, and computational simulations were conducted

⁴³See Grinols and Lin (1997, 2006) for the North-South growth model where Northern innovations are tailored to either Northern or Southern consumers.

to demonstrate the quantitative aspects of the world's socially optimal patent length. From the simulations, extending the patent length can enhance the pace of innovation, but this is accompanied by an increase in the hazard rate facing incumbent patented firms, on one hand, and by an increase in the economy's R&D intensity via increased wastes of resources arising from research congestion, or patent race, or the like, on the other hand. This paper has therefore demonstrated that extending the patent length can incur a larger opportunity cost to society in terms of the effective cost of capital and the entire economy's research intensity, as compared to previous work on patent policy, innovation, and aggregate welfare.

Applying the golden section search algorithm, the present study computed the socially optimal patent length for the calibrated model, despite that patent is always not Pareto efficient due the monopoly distortions. Its computational simulations demonstrated how the optimal patent length for a global economy is subject to the innovation-caused hazard rate, the prevalence of research congestion, and the aggregate welfare importance of product differentiation. These computational analysis and findings can be a valuable supplement to the patent policy literature, and highlight that WTO's effort to promote a uniform patent length of twenty years for the world may be debatable.

The present paper is concentrated in patent length, as there is little work attempting to quantify an economy's optimal patent length in a dynamic general equilibrium framework. It does not address the optimization of patent breadth (or patent scope), which also have effects on a patented firm's monopoly power and an economy's aggregate innovation pace. A recent work by Chu (2011) studies the welfare cost of one-size-fit-all *patent breadth* protection under an *infinite patent-length* system, using a two-sector quality-ladder growth model. From Chu's work, differentiating patent breadth protection across sectors can generate welfare gains. Certainly, if the model of the present paper extends to allow for multiple differentiated-goods sectors as in the real world, the one-size-fit-all *patent length* protection must be inferior to optimizing individual patent lengths to sectoral characteristics.

Acknowledgments

The author wishes to thank two anonymous referees for constructive comments and suggestions, and gratefully acknowledges that the paper was funded by the BarclaysAmerican Summer Research

Award of 2010 from the Belk College of Business, University of North Carolina at Charlotte. An earlier version of this paper was circulated under the title “Innovation, Creative Destruction, and Optimal Patent Life” and was presented at the 17th International Conference on Computing in Economics and Finance, June 29 - July 1, 2011, San Francisco, California, U.S.A.

Appendix

A Derivation of steady-state equilibrium condition

Note that the equilibrium value of a newly established patented firm equals $v[t, t] = v[t] = \frac{\theta w[t]}{(1-\phi g[t])n[t]}$ for which the present time t is the same as the vintage of the patent (i.e., $\tau = t$). Any firm with monopoly power earns the profit flow $\pi[t] = (\alpha - 1)w[t] \cdot x_p^M[t]$ at a time. Using (19) to replace $x_p^M[t]$, the steady-state profit flow is derived as follows

$$\pi[t] = \left(\frac{(\alpha - 1)w}{(\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon)} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \left(\frac{1}{n[t]} \right) \quad (\text{A.1})$$

where the second line holds due to $g[t] = g$, $\zeta_p^M[t] = \zeta_p^M$, and $w[t] = w$ in steady state. Using (A.1) and (11), the steady-state value of a vintage- τ patent at time t is given by

$$\begin{aligned} v[t, \tau] &= \left(\frac{(\alpha - 1)w}{(\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon)} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \int_t^{\tau+l} e^{-\int_t^s (\rho+h)d\omega} \frac{1}{n[s]} ds \\ &= \frac{1}{n[t]} \left(\frac{(\alpha - 1)w}{(\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon)} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \int_t^{\tau+l} e^{-\int_t^s (\rho+h+g)d\omega} ds \\ &= \frac{1}{n[t]} (\alpha - 1)w \left(\frac{1}{\zeta_p^M + (1 - \zeta_p^M)\alpha^\varepsilon} \right) \left(L - \frac{\theta g}{1 - \phi g} \right) \left(\frac{1 - e^{-(\rho+h+g)(\tau+l-t)}}{r + h + g} \right) \quad (\text{A.2}) \end{aligned}$$

Letting $\tau = t$ in (A.2), imposing boundary condition (13a) on $v[t, t]$ and eliminating the w term and the n term on each side of the resulting equation, we obtain (26a) in the text.

B Derivation of lifetime Utility Components

Noting that $E = 1$, $C = E/P$, $P = w \cdot n^{-\frac{1}{\varepsilon-1}} \cdot \varphi$ from (20a), and $w = E / \left[\xi \cdot \left(L - \frac{\theta g}{1-\phi g} \right) \right]$ from (24), we can derive the lifetime utility components as follows:

$$\begin{aligned}
 U &= \int_0^{\infty} e^{-\rho t} \log[n[t]^{-\beta/(\varepsilon-1)}] dt + \int_0^{\infty} e^{-\rho t} \log \frac{E[t]}{P[t]} dt \\
 &= \int_0^{\infty} e^{-\rho t} \log[n[t]^{-\beta/(\varepsilon-1)}] dt - \int_0^{\infty} e^{-\rho t} \log[w[t] \cdot n[t]^{-\frac{1}{\varepsilon-1}} \cdot \varphi[t]] dt \\
 &= \int_0^{\infty} e^{-\rho t} \log[n[t]^{(1-\beta)/(\varepsilon-1)}] dt - \int_0^{\infty} e^{-\rho t} \log[\varphi[t]] dt + \int_0^{\infty} e^{-\rho t} \log\left[L - \frac{\theta g[t]}{1-\phi g[t]}\right] dt + \int_0^{\infty} e^{-\rho t} \log[\xi[t]] dt
 \end{aligned}$$

In steady state, $g[t] = g$, $\zeta_p^M[t] = \zeta_p^M$, $n[t] = n[0]e^{gt}$, $\varphi[t] = \varphi$ from (20a) and $\xi[t] = \xi$ in terms of (25). Noting these, the above lifetime utility components can lead to (29a)-(29d).

References

- [1] Aghion, P., and Howitt, P., 1992. A model of growth through creative destruction. *Econometrica*, 60, 323–351.
- [2] Aghion, P., and P. Howitt. 1998. *Endogenous Growth Theory*. Cambridge, Mass.: The MIT Press.
- [3] Basu, Susanto. 1996. Procyclical Productivity: Increasing Returns or Cyclical Utilization? *Quarterly Journal of Economics*, 111, 709-751.
- [4] Chou, Chien-Fu & Oz Shy 1991. New Product Development and the Optimal Duration of Patents. *Southern Economic Journal*, 57, 811-821.
- [5] Chu, A. C. 2010. Effects of Patent Length on R&D: A Quantitative DGE Analysis. *Journal of Economics*, 99, 117-140.
- [6] Chu, A. C. 2011. The welfare cost of one-size-fits-all patent protection. *Journal of Economic Dynamics & Control*, 35, 876–890

- [7] Cozzi, G. 2003. Innovating on Innovations. *Rivista di Politica Economica*, 93, 3-29.
- [8] Cozzi, G., and Luca Spinesi. 2004. Information Transmission and the Bounds to Growth. *The B.E. Journal of Economic Analysis & Policy*, 4, 1-17.
- [9] Daiwan, I. and D. Rodrik. 1991. Patents, appropriate technology, and North-South trade. *Journal of International Economics*, 30, 27-47.
- [10] Davis, L. S., and F. Şener. 2012. Private patent protection in the theory of Schumpeterian growth. *European Economic Review*, 56, 1446-1460.
- [11] DeBrock, Lawrence M. 1985. Market Structure, Innovation and Optimal Patent Life. *Journal of Law and Economics*, 28, 223-244.
- [12] Dernis, H. and M. Khan. 2004. Triadic Patent Families Methodology. STI Working Paper Series, OECD.
- [13] Dinopoulos, E. and C. Syropoulos. 2007. Rent protection as a barrier to innovation and growth. *Economic Theory*, 32, 309-332.
- [14] Eicher, T., and C. García-Peñalosa. 2008. Endogenous strength of intellectual property rights: Implications for economic development and growth. *European Economic Review*, 52, 237–258.
- [15] Futagami, K. & T. Iwaisako. 2007. Dynamic analysis of patent policy in an endogenous growth model. *Journal of Economic Theory*, 132, 306 - 334.
- [16] Furukawa, Yuichi. 2007. The protection of intellectual property rights and endogenous growth: Is stronger always better? *Journal of Economic Dynamics and Control*, 31, 3644–3670.
- [17] Gilbert, R., & Shapiro, C. 1990. Optimal Patent Length and Breadth. *Rand Journal of Economics* 21, 106–112.
- [18] Grieben, W. and Fuat Şener. 2009. Globalization, rent protection institutions, and going alone in freeing trade. *European Economic Review*, 53, 1042-1065.
- [19] Grinols, L. Earl & H. C. Lin. 1997. Asymmetric intellectual property rights protection and North-South welfare. Office of Research Working Paper No. 98-0106, University of Illinois.

- [20] Grinols, E. L. & H. C. Lin, 2006. Global patent protection: channels of North & South welfare gain. *Journal of Economic Dynamics & Control* 30, 205-227.
- [21] Grinols, E. L. & H. C. Lin, 2011. Patent replacement and welfare gains. *Journal of Economic Dynamics and Control*, 35, 1586-1604.
- [22] Grossman, G. M., and E. Helpman. 1992. *Innovation and Growth in the Global Economy*. Cambridge, Mass.: The MIT Press.
- [23] Grossman, G., and Lai, E., 2004. International protection of intellectual property. *American Economic Review*, 94, 1635–1653.
- [24] Helpman, E., 1993. Innovation, Imitation, and Intellectual Property Rights. *Econometrica*, 61, 1247-90.
- [25] Iwaisako, T., H. Tanaka, and K. Futagami. 2011. A Welfare analysis of global patent protection in a model with endogenous innovation and foreign direct investment. *European Economic Review*, 55, 1137-1151.
- [26] Iwaisako, T. & K. Futagami. 2003. Patent policy in an endogenous growth model. *Journal of Economics* 78, 239–258.
- [27] Judd, K. L., 1985. On the performance of patents. *Econometrica*, 53, 567-85.
- [28] Jones, Charles I. 1999. Growth: with or without scale effects? *American Economic Review*, 89, 139-144.
- [29] Jones, C. I., 1995. R&D-based models of economic growth. *Journal of Political Economy*, 103, 759–784.
- [30] Jones, C. I., and J. C. Williams. 2000. Too much of a good thing? The economics of investment in R&D. *Journal of Economic Growth* , 5, 65 - 85.
- [31] Kiusalaas, Jaan. 2010. *Numerical Methods in Engineering with Python*, Second Edition. New York, NY: Cambridge University Press.
- [32] Lin, Hwan C. 2010. Technology Diffusion and Global Welfare Effects: Imitative R&D vs. South-Bound FDI. *Structural Change and Economic Dynamics*. Vol. 21, Issue 4, 231-247.

- [33] Lin, Hwan C. 2002. Shall the Northern Optimal R&D Subsidy Rate Inversely Respond to Southern Intellectual Property Protection? *Southern Economic Journal*, 69, 381-397.
- [34] Nordhaus, William D. 1969. *Invention, Growth and Welfare*. Cambridge, Mass.: MIT Press.
- [35] Norrbin, Stefan C. 1993. The Relationship Between Price and Marginal Cost in U.S. Industry: A Contradiction. *Journal of Political Economy*, 101, 1149-1164.
- [36] O'Donoghue, T. & J. Zweimüller. 2004. Patents in a Model of Endogenous Growth. *Journal of Economic Growth*, 9, 81-123.
- [37] Romer, Paul M. 1990. Endogenous Technological Change. *Journal of Political Economy*, 98, 71-102.
- [38] Scherer, F. M., 1972. Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation. *American Economic Review*, 62, 422-427.
- [39] Segerstrom, P., Anant, & T., Dinopoulos, E., 1990. A Schumpeterian model of the product life cycle. *American Economic Review*, 80, 1077-1092.
- [40] Wright, B. D. 1983. The Economics of Invention Incentives: Patents, Prizes and Research Contracts. *American Economic Review*, 73, 691-707.